

PCT

01/50940

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :  C12N 15/00, C07K 14/47, G01N 33/53		A2	(11) International Publication Number: <b>WO 00/52151</b>  (43) International Publication Date: 8 September 2000 (08.09.00)
<p>(21) International Application Number: PCT/US00/05621</p> <p>(22) International Filing Date: 3 March 2000 (03.03.00)</p> <p>(30) Priority Data: 60/123,117 5 March 1999 (05.03.99) US</p> <p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 60/123,117 (CIP) Filed on 5 March 1999 (05.03.99)</p> <p>(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). LU, Dyung, Aina, M. [US/US]; 55 Park</p>		<p>Belmont Place, San Jose, CA 95136 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US).</p> <p>(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(81) Designated States: AB, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>	
<p>(54) Title: HUMAN SECRETORY PROTEINS</p> <p>(57) Abstract</p> <p>The invention provides human secretory proteins (HSECP) and polynucleotides which identify and encode HSECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HSECP.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Malí	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## HUMAN SECRETORY PROTEINS

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human secretory proteins 5 and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders.

### BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is 10 mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. 15 Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Secreted proteins are often synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal 20 peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include receptors, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, neuropeptides, vasomediators, ion channels, transporters/pumps, and proteases. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557- 25 560, 582-592.)

G-protein coupled receptors (GPCRs) comprise a superfamily of integral membrane proteins which transduce extracellular signals. Not all GPCRs contain N-terminal signal peptides. GPCRs include receptors for biogenic amines such as dopamine, epinephrine, histamine, glutamate (metabotropic-type), acetylcholine (muscarinic-type), and serotonin; for lipid mediators of 30 inflammation such as prostaglandins, platelet activating factor, and leukotrienes; for peptide hormones such as calcitonin, C5a anaphylatoxin, follicle stimulating hormone, gonadotropin releasing hormone, neurokinin, oxytocin, and thrombin; and for sensory signal mediators such as retinal photopigments and olfactory stimulatory molecules. The structure of these highly conserved receptors consists of seven hydrophobic transmembrane regions, cysteine disulfide bridges between 35 the second and third extracellular loops, an extracellular N-terminus, and a cytoplasmic C-terminus.

The N-terminus interacts with ligands, the disulfide bridges interact with agonists and antagonists, and the large third intracellular loop interacts with G proteins to activate second messengers such as cyclic AMP, phospholipase C, inositol triphosphate, or ion channels. (Reviewed in Watson, S. and Arkinstall, S. (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp. 2-6; and Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego, CA, pp. 162-176.)

Other types of receptors include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI).

(Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) *Kidney Int.* 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) *BioEssays* 19:47-55.)

Cytokines are secreted by hematopoietic cells in response to injury or infection. Interleukins, neurotrophins, growth factors, interferons, and chemokines all define cytokine families that work in

conjunction with cellular receptors to regulate cell proliferation and differentiation. In addition, cytokines effect activities such as leukocyte migration and function, hematopoietic cell proliferation, temperature regulation, acute response to infection, tissue remodeling, and apoptosis.

Chemokines, in particular, are small chemoattractant cytokines involved in inflammation,

5 leukocyte proliferation and migration, angiogenesis and angiostasis, regulation of hematopoiesis, HIV infectivity, and stimulation of cytokine secretion. Chemokines generally contain 70-100 amino acids and are subdivided into four subfamilies based on the presence of conserved cysteine-based motifs.

(Callard, R. and Gearing, A. (1994) The Cytokine Facts Book, Academic Press, New York, NY, pp. 181-190, 210-213, 223-227.)

10 Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with MPs for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There  
15 are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs).

Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-  
20 lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin which function as hormones to regulate cellular functions other than proliferation.

25 Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition,  
30 growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes.

(Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

Proteolytic enzymes or proteases either activate or deactivate proteins by hydrolyzing peptide bonds. Proteases are found in the cytosol, in membrane-bound compartments, and in the extracellular space. The major families are the zinc, serine, cysteine, thiol, and carboxyl proteases.

35 Ion channels, ion pumps, and transport proteins mediate the transport of molecules across

cellular membranes. Transport can occur by a passive, concentration-dependent mechanism or can be linked to an energy source such as ATP hydrolysis. Symporters and antiporters transport ions and small molecules such as amino acids, glucose, and drugs. Symporters transport molecules and ions unidirectionally, and antiporters transport molecules and ions bidirectionally. Transporter

5 superfamilies include facilitative transporters and active ATP-binding cassette transporters which are involved in multiple-drug resistance and the targeting of antigenic peptides to MHC Class I molecules. These transporters bind to a specific ion or other molecule and undergo a conformational change in order to transfer the ion or molecule across the membrane. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 523-546.)

10 Ion channels are formed by transmembrane proteins which create a lined passageway across the membrane through which water and ions, such as  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ , and  $\text{Cl}^-$ , enter and exit the cell. For example, chloride channels are involved in the regulation of the membrane electric potential as well as absorption and secretion of ions across the membrane. Chloride channels also regulate the internal pH of membrane-bound organelles.

15 Ion pumps are ATPases which actively maintain membrane gradients. Ion pumps are classified as P, V, or F according to their structure and function. All have one or more binding sites for ATP in their cytosolic domains. The P-class ion pumps include  $\text{Ca}^{2+}$  ATPase and  $\text{Na}^+/\text{K}^+$  ATPase and function in transporting  $\text{H}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ , and  $\text{Ca}^{2+}$  ions. P-class pumps consist of two  $\alpha$  and two  $\beta$  transmembrane subunits. The V- and F-class ion pumps have similar structures but transport only  $\text{H}^+$ .  
20 F class  $\text{H}^+$  pumps mediate transport across the membranes of mitochondria and chloroplasts, while V-class  $\text{H}^+$  pumps regulate acidity inside lysosomes, endosomes, and plant vacuoles.

A family of structurally related intrinsic membrane proteins known as facilitative glucose transporters catalyze the movement of glucose and other selected sugars across the plasma membrane. The proteins in this family contain a highly conserved, large transmembrane domain comprised of 12 25  $\alpha$ -helices, and several weakly conserved, cytoplasmic and exoplasmic domains. (Pessin, J. E., and Bell, G.I. (1992) *Annu. Rev. Physiol.* 54:911-930.)

Amino acid transport is mediated by  $\text{Na}^+$  dependent amino acid transporters. These transporters are involved in gastrointestinal and renal uptake of dietary and cellular amino acids and in neuronal reuptake of neurotransmitters. Transport of cationic amino acids is mediated by the 30 system y<sup>+</sup> family and the cationic amino acid transporter (CAT) family. Members of the CAT family share a high degree of sequence homology, and each contains 12-14 putative transmembrane domains. (Ito, K. and Groudine, M. (1997) *J. Biol. Chem.* 272:26780-26786.)

Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical 35 compositions and mechanisms of action, hormones can be grouped into two categories. One category

includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes 5 hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific 10 American Books Inc., New York, NY, pp. 856-864.)

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neuropeptid Y, neuropeptid N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, 15 vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C. R. et al. (1985) Endocrine 20 Physiology, Oxford University Press, New York, NY, pp. 57-62.)

The discovery of new human secretory proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders.

25

#### SUMMARY OF THE INVENTION

The invention features purified polypeptides, human secretory proteins, referred to collectively as "HSECP" and individually as "HSECP-1," "HSECP-2," "HSECP-3," "HSECP-4," "HSECP-5," "HSECP-6," "HSECP-7," "HSECP-8," "HSECP-9," "HSECP-10," "HSECP-11," 30 "HSECP-12," "HSECP-13," "HSECP-14," "HSECP-15," "HSECP-16," "HSECP-17," "HSECP-18," "HSECP-19," "HSECP-20," "HSECP-21" and "HSECP-22." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active 35 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an

immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-22.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising

5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. In one

10 alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:23-44.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group

15 consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

20 The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino

25 acid sequence selected from the group consisting of SEQ ID NO:1-22. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

30 Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID

35 NO:1-22.

The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, c) a polynucleotide sequence complementary to a), 5 or d) a polynucleotide sequence complementary to b). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). The method comprises a) 10 hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

20 The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID 25 NO:1-22, and a pharmaceutically acceptable excipient. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HSECP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ 35 ID NO:1-22. The method comprises a) exposing a sample comprising the polypeptide to a

compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional HSECP,

5 comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22,

10 c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by

15 the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional HSECP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method for screening a compound for effectiveness in

20 altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:23-44, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

25

#### BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HSECP.

Table 2 shows features of each polypeptide sequence, including predicted signal peptides and

30 other motifs, and methods, algorithms, and searchable databases used for analysis of HSECP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

35 Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones

encoding HSECP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HSECP, along with applicable descriptions, references, and threshold parameters.

5

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which 10 will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so 15 forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now 20 described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

## DEFINITIONS

25 "HSECP" refers to the amino acid sequences of substantially purified HSECP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of HSECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other 30 compound or composition which modulates the activity of HSECP either by directly interacting with HSECP or by acting on components of the biological pathway in which HSECP participates.

An "allelic variant" is an alternative form of the gene encoding HSECP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in 35 polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to

allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding HSECP include those sequences with deletions, 5 insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HSECP or a polypeptide with at least one functional characteristic of HSECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HSECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding 10 HSECP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HSECP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HSECP is retained. For example, 15 negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

20 The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to an amino acid sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein 25 molecule.

“Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity 30 of HSECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HSECP either by directly interacting with HSECP or by acting on components of the biological pathway in which HSECP participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments 35 thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant.

Antibodies that bind HSECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired.

5 Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to 10 immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

15 The term "antisense" refers to any composition capable of base-pairing with the "sense" strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be 20 produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

25 The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HSECP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

30 The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization 35 between the nucleic acid strands. This is of particular importance in amplification reactions, which

depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HSECP or fragments of HSECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

10 "Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer 15 program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

20 "Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded 25 as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
25	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
30	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
35	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
40	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

5 A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative 10 polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of HSECP or the polynucleotide encoding HSECP which is 15 identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid 20 residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present 25 embodiments.

A fragment of SEQ ID NO:23-44 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:23-44, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:23-44 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:23-44 from related 30 polynucleotide sequences. The precise length of a fragment of SEQ ID NO:23-44 and the region of SEQ ID NO:23-44 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-22 is encoded by a fragment of SEQ ID NO:23-44. A fragment of SEQ ID NO:1-22 comprises a region of unique amino acid sequence that specifically identifies 35 SEQ ID NO:1-22. For example, a fragment of SEQ ID NO:1-22 is useful as an immunogenic peptide

for the development of antibodies that specifically recognize SEQ ID NO:1-22. The precise length of a fragment of SEQ ID NO:1-22 and the region of SEQ ID NO:1-22 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

5 The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization 10 assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to 15 one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

20 The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

25 Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 30 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

35 Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment

Search Tool (BLAST) (Altschul, S.F. et al. (1990) *J. Mol. Biol.* 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

15            *Matrix: BLOSUM62*  
          *Reward for match: 1*  
          *Penalty for mismatch: -2*  
          *Open Gap: 5 and Extension Gap: 2 penalties*  
          *Gap x drop-off: 50*  
          *Expect: 10*  
          *Word Size: 11*  
          *Filter: on*  
20            Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length  
25            supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

30            Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

35            The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the

site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of 5 polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise 10 comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Open Gap: 11 and Extension Gap: 1 penalties*

*Gap x drop-off: 50*

15 *Expect: 10*

*Word Size: 3*

*Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for 20 example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

25 "Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

30 The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized 35 after the "washing" step(s). The washing step(s) is particularly important in determining the

stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C<sub>0</sub>t or K<sub>0</sub>t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively. "Immune response" can refer to conditions associated with inflammation, trauma, immune

disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of HSECP which is

5 capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of HSECP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

10 The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HSECP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HSECP.

15 The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

20 "Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

25 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

30 "Probe" refers to nucleic acid sequences encoding HSECP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target

35 polynucleotide by complementary base-pairing. The primer may then be extended along the target

DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also 5 be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for 10 example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose 15 such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer 20 selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for 25 Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping 30 Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization 35 technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to

identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence.

5 This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter

10 sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a *vaccinia* virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

15 An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

20 The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HSECP, or fragments thereof, or HSECP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

25 The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

30 The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

35 A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

5 "Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral  
10 infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to  
15 animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in  
20 vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention  
25 into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-  
30 1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of  
35 polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding

polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene 5 between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having 10 at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

## 15 THE INVENTION

The invention is based on the discovery of new human secretory proteins (HSECP), the polynucleotides encoding HSECP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders.

20 Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HSECP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HSECP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their 25 corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each HSECP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: 30 column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; and column 6 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 6 were used to characterize each polypeptide through sequence homology and protein motifs. In column 5, the first line of each cell lists the amino acid 35 residues comprising predicted signal peptide sequences located at the amino terminus of each

HSECP. Additional identifying motifs or signatures, such as a somatomedin B signature in SEQ ID NO:16 and seven putative transmembrane domains in SEQ ID NO:18, are also listed in column 5.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HSECP. The first column of Table 3 lists the 5 nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:23-44 and to distinguish between SEQ ID NO:23-44 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express HSECP as a fraction of total tissues expressing HSECP.

10 Column 4 lists diseases, disorders, or conditions associated with those tissues expressing HSECP as a fraction of total tissues expressing HSECP. Column 5 lists the vectors used to subclone each cDNA library. In particular, three out of four cDNA libraries which express SEQ ID NO:23 are derived from cartilage and synovia associated with joint inflammation, and four out of five cDNA libraries which express SEQ ID NO:29 are derived from intestinal tissue. Furthermore, about half of the

15 15 cDNA libraries expressing SEQ ID NO:34 are associated with inflammation or the hematopoietic/immune system. Likewise, about half of the cDNA libraries expressing SEQ ID NO:35 are associated with inflammation or the hematopoietic/immune system, and in particular, with inflammation of the joints. In addition, 82% of the cDNA libraries expressing SEQ ID NO:37 are derived from tissues of the nervous system. Finally, expression of SEQ ID NO:39 is detected solely

20 20 in a subtracted prostate tumor cDNA library, and expression of SEQ ID NO:43 is detected only in two cDNA libraries derived from heart tissue.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HSECP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 25 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses HSECP variants. A preferred HSECP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HSECP amino acid sequence, and which contains at least one functional or structural characteristic of HSECP.

30 The invention also encompasses polynucleotides which encode HSECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:23-44, which encodes HSECP. The polynucleotide sequences of SEQ ID NO:23-44, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the 35 sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding HSECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HSECP. A particular aspect of the invention encompasses a variant of a 5 polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:23-44 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:23-44. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HSECP.

10 It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HSECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These 15 combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HSECP, and all such variations are to be considered as being specifically disclosed.

20 Although nucleotide sequences which encode HSECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HSECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HSECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with 25 which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HSECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

30 The invention also encompasses production of DNA sequences which encode HSECP and HSECP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HSECP or any fragment thereof.

35 Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:23-44 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and

G33S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of 5 the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is 10 automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a 15 variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HSECP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, 20 such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown 25 sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme 30 digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed 35 using commercially available software, such as OLIGO 4.06 Primer Analysis software (National

Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been

5 size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze

10 the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process  
15 from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HSECP may be cloned in recombinant DNA molecules that direct expression of

20 HSECP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HSECP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HSECP-encoding sequences for a variety of purposes including, but  
25 not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

30 The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) *Nat. Biotechnol.* 17:793-797; Christians, F.C. et al. (1999) *Nat. Biotechnol.* 17:259-264; and Crameri, A. et al. (1996) *Nat. Biotechnol.* 14:315-319) to alter or improve the biological properties of HSECP, such as its biological or enzymatic activity or its ability  
35 to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene

variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial"

5 breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable

10 manner.

In another embodiment, sequences encoding HSECP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.)

15 Alternatively, HSECP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HSECP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

20 The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) *Proteins, Structures and Molecular Properties*, WH Freeman, New York NY.)

25 In order to express a biologically active HSECP, the nucleotide sequences encoding HSECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in

30 polynucleotide sequences encoding HSECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HSECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HSECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional

35 transcriptional or translational control signals may be needed. However, in cases where only coding

sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used.

5 (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HSECP and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HSECP. These include, but are not limited to, microorganisms such as bacteria transformed 15 with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

20 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HSECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HSECP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HSECP into the vector's multiple 25 cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of HSECP are needed, e.g. for the production of 30 antibodies, vectors which direct high level expression of HSECP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

35 Yeast expression systems may be used for production of HSECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable

integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of HSECP. Transcription of sequences

5 encoding HSECP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.)

10 These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HSECP may be ligated into 15 an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HSECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 25: 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HSECP in cell lines is preferred. For example, sequences encoding HSECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. 30 Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

35 Any number of selection systems may be used to recover transformed cell lines. These

include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) *Cell* 11:223-232; Lowy, I. et al. (1980) *Cell* 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to 5 methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbere-Garapin, F. et al. (1981) *J. Mol. Biol.* 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

15 Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HSECP is inserted within a marker gene sequence, transformed cells containing sequences encoding HSECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HSECP under the control of a 20 single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HSECP and that express HSECP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR 25 amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

30 Immunological methods for detecting and measuring the expression of HSECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HSECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and 35 Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana

Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HSECP

5 include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HSECP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety  
10 of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HSECP may be cultured under  
15 conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HSECP may be designed to contain signal sequences which direct secretion of HSECP through a prokaryotic or eukaryotic cell membrane.

20 In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity.

25 Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid  
30 sequences encoding HSECP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HSECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HSECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available  
35 affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST),

maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity

5 purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HSECP encoding sequence and the heterologous protein sequence, so that HSECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10).

10 A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HSECP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the

15 T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example,  $^{35}\text{S}$ -methionine.

Fragments of HSECP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, *supra*, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be

20 achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HSECP may be synthesized separately and then combined to produce the full length molecule.

## Therapeutics

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HSECP and human secretory proteins. In addition, the expression of HSECP is

25 closely associated with cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders. Therefore, HSECP appears to play a role in cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders. In the treatment of disorders associated with increased HSECP expression or activity, it is desirable to decrease the expression or activity of HSECP. In the treatment of disorders associated with decreased HSECP expression or activity, it is desirable to

30 increase the expression or activity of HSECP.

Therefore, in one embodiment, HSECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in

35 particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall

bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, 5 atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's, 10 thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal 15 circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, 20 pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic 25 encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha-1 antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatitis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a 30 cardiovascular disorder, and in particular, a disorder of the heart such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus 35 erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart

disease, congenital heart disease, and complications of cardiac transplantation; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural

5 muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the

10 nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic,

15 endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder.

In another embodiment, a vector capable of expressing HSECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HSECP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HSECP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HSECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSECP. Examples of such disorders include, but are not limited to, those cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders described above. In one aspect, an antibody which specifically binds HSECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express HSECP.

35 In an additional embodiment, a vector expressing the complement of the polynucleotide

encoding HSECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSECP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HSECP may be produced using methods which are generally known in the art. In particular, purified HSECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HSECP. Antibodies to HSECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HSECP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HSECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HSECP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HSECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and

Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc.

5 Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HSECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g.,

10 Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

15 Antibody fragments which contain specific binding sites for HSECP may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D.

20 et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HSECP and its 25 specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HSECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HSECP. Affinity is expressed as an 30 association constant,  $K_a$ , which is defined as the molar concentration of HSECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HSECP epitopes, represents the average affinity, or avidity, of the antibodies for HSECP. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific 35 for a particular HSECP epitope, represents a true measure of affinity. High-affinity antibody

preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the HSECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HSECP,

5 preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For 10 example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HSECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

15 In another embodiment of the invention, the polynucleotides encoding HSECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HSECP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HSECP. Thus, complementary molecules or 20 fragments may be used to modulate HSECP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HSECP.

25 Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HSECP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

30 Genes encoding HSECP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HSECP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of 35 the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HSECP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed.

5 Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco 10 NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, 15 engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HSECP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, 20 corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared 25 by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HSECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA 30 constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase 35 linkages within the backbone of the molecule. This concept is inherent in the production of PNAs

and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

5 Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) *Nat.*

10 *Biotechnol.* 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical 15 or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HSECP, antibodies to HSECP, and mimetics, agonists, antagonists, or inhibitors of HSECP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, 20 but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any 25 number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceuticals-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's 30 Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceuticals-acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

35 Pharmaceutical preparations for oral use can be obtained through combining active

compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose,

5 hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragée cores may be used in conjunction with suitable coatings, such as concentrated sugar 10 solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of 15 gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

20 Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily 25 injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

30 For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

35 The pharmaceutical composition may be provided as a salt and can be formed with many

acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a 5 pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HSECP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein 10 the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of 15 administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HSECP or fragments thereof, antibodies of HSECP, and agonists, antagonists or inhibitors of HSECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined 20 by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population) or LD<sub>50</sub> (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD<sub>50</sub>/ED<sub>50</sub> ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and 25 animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED<sub>50</sub> with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the 30 subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 35 days, every week, or biweekly depending on the half-life and clearance rate of the particular

formulation.

Normal dosage amounts may vary from about 0.1  $\mu$ g to 100,000  $\mu$ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

5 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

#### DIAGNOSTICS

In another embodiment, antibodies which specifically bind HSECP may be used for the

10 diagnosis of disorders characterized by expression of HSECP, or in assays to monitor patients being treated with HSECP or agonists, antagonists, or inhibitors of HSECP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HSECP include methods which utilize the antibody and a label to detect HSECP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or 15 without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HSECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HSECP expression.

20 Normal or standard values for HSECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to HSECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HSECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. 25 Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HSECP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HSECP may be correlated 30 with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HSECP, and to monitor regulation of HSECP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HSECP or closely related molecules may be used to identify nucleic acid sequences which encode HSECP. The specificity of the probe, whether it is 35 made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a

conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HSECP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the HSECP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:23-44 or from genomic sequences including promoters, enhancers, and introns of the HSECP gene.

Means for producing specific hybridization probes for DNAs encoding HSECP include the cloning of polynucleotide sequences encoding HSECP or HSECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HSECP may be used for the diagnosis of disorders associated with expression of HSECP. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea,

emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis,

5 Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha - antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein

10 obstruction and thrombosis, centrilobular necrosis, peliosis hepatitis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a cardiovascular disorder, and in particular, a disorder of the heart such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease,

15 degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; and a neurological

20 disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess,

25 suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central

30 nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia,

35 catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic

neuralgia, and Tourette's disorder. The polynucleotide sequences encoding HSECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HSECP expression. Such qualitative or quantitative methods are well

5 known in the art.

In a particular aspect, the nucleotide sequences encoding HSECP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HSECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a 10 suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HSECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to 15 monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HSECP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HSECP, under conditions suitable for hybridization or 20 amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

25 Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

30 With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development 35 or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HSECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HSECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding

5 HSECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HSECP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from 10 standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

15 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and 20 monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-25 2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HSECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial 30 chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, 35 pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the

Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding HSECP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene

5 sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not

10 known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti,

15 R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HSECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HSECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HSECP, or fragments thereof, and washed. Bound HSECP is then detected by methods well known in the art. Purified HSECP can also be coated directly onto plates for use in the aforementioned drug screening techniques.

Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HSECP specifically compete with a test compound for binding HSECP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HSECP.

35 In additional embodiments, the nucleotide sequences which encode HSECP may be used in

any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding 5 description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

Without further elaboration, it is believed that one skilled in the art can, using the preceding 10 description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/123,117, are hereby expressly incorporated by reference.

15

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a 20 monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA 25 purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA 30 libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the 35 appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-

1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSCAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

15 The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire 20 annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying 25 against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosise, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

30 The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:23-44. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Northern Analysis

35 Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs

from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This 5 analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

10 The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

15 The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HSECP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, 20 cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### V. Extension of HSECP Encoding Polynucleotides

25 The full length nucleic acid sequences of SEQ ID NO:23-44 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68 °C to about 72 °C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

35 High fidelity amplification was obtained by PCR using methods well known in the art. PCR

was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer

5 pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

10 The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the 15 concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and 20 sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site 25 overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following 30 parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing 35 primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM

BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:23-44 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

5 **VI. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:23-44 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06  
10 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech).  
An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based  
15 hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature  
20 under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

**VII. Microarrays**

A chemical coupling procedure and an ink jet device can be used to synthesize array  
25 elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and  
30 patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software  
35 well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or

fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. 5 (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

### VIII. Complementary Polynucleotides

Sequences complementary to the HSECP-encoding sequences, or any parts thereof, are used 10 to detect, decrease, or inhibit expression of naturally occurring HSECP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HSECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence 15 and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HSECP-encoding transcript.

### IX. Expression of HSECP

Expression and purification of HSECP is achieved using bacterial or virus-based expression 20 systems. For expression of HSECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). 25 Antibiotic resistant bacteria express HSECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HSECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HSECP by either homologous recombination or bacterial-mediated 30 transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 35 7:1937-1945.)

In most expression systems, HSECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HSECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified HSECP obtained by these methods can be used directly in the following activity assay.

#### X. Demonstration of HSECP Activity

An assay for HSECP activity measures the expression of HSECP on the cell surface. cDNA encoding HSECP is subcloned into an appropriate mammalian expression vector suitable for high levels of cDNA expression. The resulting construct is transfected into a nonhuman cell line such as NIH3T3. Cell surface proteins are labeled with biotin using methods known in the art. Immunoprecipitations are performed using HSECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of HSECP expressed on the cell surface.

Alternatively, an assay for HSECP activity measures the amount of HSECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles. Immunoprecipitations from fractionated and total cell lysates are performed using HSECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The concentration of HSECP in secretory organelles relative to HSECP in total cell lysate is proportional to the amount of HSECP in transit through the secretory pathway.

#### XI. Functional Assays

HSECP function is assessed by expressing the sequences encoding HSECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1

plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein 5 provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects 10 and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as 15 measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HSECP on gene expression can be assessed using highly purified 20 populations of cells transfected with sequences encoding HSECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions 25 of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HSECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

## XII. Production of HSECP Specific Antibodies

HSECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to 30 immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HSECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

35 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A

peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-

5 HSECP activity by, for example, binding the peptide or HSECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

### XIII. Purification of Naturally Occurring HSECP Using Specific Antibodies

Naturally occurring or recombinant HSECP is substantially purified by immunoaffinity chromatography using antibodies specific for HSECP. An immunoaffinity column is constructed by

10 covalently coupling anti-HSECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HSECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HSECP (e.g., high ionic strength

15 buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HSECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HSECP is collected.

### XIV. Identification of Molecules Which Interact with HSECP

HSECP, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter

20 reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HSECP, washed, and any wells with labeled HSECP complex are assayed. Data obtained using different concentrations of HSECP are used to calculate values for the number, affinity, and association of HSECP with the candidate molecules.

25 Alternatively, molecules interacting with HSECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

Various modifications and variations of the described methods and systems of the invention

30 will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the

35 scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments	
1	23	078811	SYNORAB01	078811H1 (SYNORAB01), 078811T6 (SYNORAB01),	078811R6 (SYNORAB01),
2	24	371156	LUNGNOT02	077182F1 (SYNORAB01), 289599H1 (TMLR3DT01), 523415H1 (MMLR2DT01), 3772179F6 (BRSTNOT25),	077182R1 (SYNORAB01), 371156H1 (LUNGNOT02), 3438708F6 (PENCNOT06), 3808004H1 (CONTUT01)
3	25	584050	PROSNOT02	584050H1 (PROSNOT02), 1739666R6 (HIPONON01), 2305379R6 (NGANNOT01), 4070575H1 (KIDNNOT26), SZA01719F1, SASA02714F1	1726563T6 (PROSNOT14), 1856214F6 (PROSNOT18), 2681374F6 (SINIUCT01), 5274539H1 (OVARDIN02),
4	26	863808	BRAITUT03	851821H1 (NGANNOT01), 863808T1 (BRAITUT03),	863808H1 (BRAITUT03), 2735728F6 (OVARNOT09)
5	27	978433	BRSTNOT02	054823R1 (FIBRNOT01), 978433R1 (BRSTNOT02), 2503122H1 (CONUTU01), 3411659H1 (BRSTTUS08)	978433H1 (BRSTNOT02), 1867687T6 (SKINBIT01), 2532586H1 (GBLANOT02),
6	28	1655369	PROSTUT08	746013R1 (BRAITUT01), 1539790R1 (SINTTUT01), 1655369F6 (PROSTUT08), 1673290F6 (BLADNOT05), 3407992H1 (PROSTUS08), 4098012H1 (BRAITUT26)	944864R6 (RATRNOT02), 1617847F6 (BRAITUT12), 1655369H1 (PROSTUT08), 2056840R6 (BEPINOT01), 4077365H1 (PANCNOT19),

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
7	29	1703244	DUODNOT02	273878H1 (PANC01T03), 1632051H1 (COL01N019), 1703244H1 (DUODNOT02), 4176224H1 (SINT021)
8	30	1730819	BRSTTUT08	862188R1 (BRAITUT03), 1399644F1 (BRAITUT08), 1443690F1 (THYRN0T03), 1596446F1 (BRAIN0T14), 1730819F6 (BRSTTUT08), 1730819H1 (BRSTTUT08), 2304942H1 (NGANN0T01), 2868843H1 (THYRN0T10), 3395457H1 (LUNG0T28)
9	31	1757161	PITUN0T03	864626R1 (BRAITUT03), 1231577H1 (BRAITUT01), 1395593F1 (THYRN0T03), 1466766T1 (PANC01T02), 1466766T6 (PANC01T02), 1597583F6 (BRAIN0T14), 1757161H1 (PITUN0T03), 1757161R6 (PITUN0T03), 1757161T6 (PITUN0T03)
10	32	1976095	PANC01T02	864976T1 (BRAITUT03), 1976095H1 (PANC01T02), 4616101H1 (BRAYD01T01), SARB01143F1, SARB01861F1, SAJA02355F1
11	33	2169991	ENDC01T03	2169991H1 (ENDC01T03), 2313548R3 (NGANN0T01), 2727735T3 (OVART05), 3095616H1 (CERV01T03), 3186421H1 (THYMN04), 3493007H1 (ADRETT07), SASA03601F1, SBKA05288F1, SBDA00157F1
12	34	2616827	GBLAN0T01	2616827H1 (GBLAN0T01), 4637616F6 (MYEPTXT01), 5219252H1 (BRST0N0T35)
13	35	2991370	KIDNFET02	080447R1 (SYNORAB01), 190292R1 (SYNORAB01), 192279R1 (SYNORAB01), 2724873H1 (OVART05), 2991370H1 (KIDNFET02), SAVA03596F1, SAVA01525F1

Table 1 (cont.)

Polypeptide SEQ ID NO:		Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
14	36	3031062	TLYMNNOT05	1354027T3 (LUNGNOT09), 2889215F6 (LUNGFEET04), 3031062H1 (TLYMNNOT05), SBAA04235F1, SBAA00620F1, SBAA01760F1, SBAA04365F2, SBAA01589F1	
15	37	3101617	BRAINOT20	3101617H1 (BRAINOT20), 3335717F6 (BRAIFET01), SCAA04574V1, SCAA04351V1, SCAA03628V1, SCAA05459V1, SCAA01004V1	
16	38	3216178	TESTNOT07	632084R6 (KIDNNNOT05), 1993593T6 (CORPNOT02), 3216178F6 (TESTNOT07), 3216178H1 (TESTNOT07), 4914242H1 (LIVRFET05)	
17	39	3406803	PROSTUS08	3406803F6 (PROSTUS08), 3406803H1 (PROSTUS08), 3406803T6 (PROSTUS08)	
18	40	3468066	BRAIDIT01	659544H1 (BRAINOT03), 897142R1 (BRSTNOT05), 1321038F1 (BLADNOT04), 1351888F1 (LATRTUT02), 1485695F1 (CORPNOT02), 1507666F1 (LUNGNOT14), 2953291H1 (KIDNFEET01), 3468056H1 (BRAIDIT01), 4426018F6 (BRAPDIT01)	
19	41	3592862	293TF5T01	941610H1 (ADRENOT03), 1288036H1 (BRAINOT11), 1687969F6 (PROSTUT10), 3592862H1 (293TF5T01)	
20	42	3669422	KIDNTUT16	3669422F6 (KIDNTUT16), 3669422H1 (KIDNTUT16), 3669422T6 (KIDNTUT16), 5445503H1 (LNODNOT12)	
21	43	3688740	HEAANOT01	462098R6 (LATRNOT01), 3688740H1 (HEAANOT01)	
22	44	3742589	THYMNNOT08	938795R1 (CERVNOT01), 1255960T1 (MENITUT03), 1530330R1 (PANCNOT04), 3742589H1 (THYMNOT08), SBOA04142D1	

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Peptides and Other Signature Sequences	Analytical Methods
1	182	S149 T151 S13		M1-S16; M1-P18	MOTIFS SPSCAN HMM
2	125	T87 S36 T78 T111	N75	M1-S37	MOTIFS SPSCAN HMM
3	320	S21 T63 S267 S300 T164	N40	M1-G20; M1-G23	MOTIFS SPSCAN HMM
4	234	T74 S198 T210 T227 S131 T195		M1-A30; M1-G25	MOTIFS SPSCAN HMM
5	278	S64 S132 S230 T252 S7 S179	N221	M1-A65	MOTIFS SPSCAN HMM
6	136	S98 S99		M1-P54	MOTIFS SPSCAN HMM
7	109	T57		M1-T20	MOTIFS SPSCAN HMM

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Peptides and Other Signature Sequences	Potential Signal Sequences	Analytical Methods
8	262	S73 S91 S136 S86	N182	M1-A25; M1-G27	MOTIFS SPSCAN HMM	
9	384	S11 T140 S32 S185 S232 S306 S378	N50 N59 N62 N304	M1-G31 Transmembrane domains: G195-A220 L74-F91 A142-T160	MOTIFS SPSCAN HMM	
10	244	T95 S110 T208 T44 S47 S53 S69 S152 T194		M1-A23	MOTIFS SPSCAN HMM	
11	326	S138 T59 S239	N129 N237	M1-G40	MOTIFS SPSCAN HMM	
12	105	S96 T24	N94	M1-A19; M1-A23	MOTIFS SPSCAN HMM	

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Peptides and Other Signature Sequences	Analytical Methods
13	626	S188 S593 T61 T118 T144 S252 T275 T410 S423 S501 T506 S524 T536 S550 S28 T146 S336 T467 S583 T588		M1-G27 RGD tripeptide: R271-D273	MOTIFS SPSCAN HMM
14	296	S205 T295 S109 T165 T214 S244 S73 S225 Y236	N40 N53 N204 N281	M1-G23	MOTIFS SPSCAN HMM
15	249	S7 T123 S233 S237 T151	N100	M1-G23; M1-A22 Transmembrane domains: T175-A193 P97-F120	MOTIFS SPSCAN HMM
16	124	S62 S109		M1-G37 Somatomedin B signature: L36-S105	MOTIFS SPSCAN HMM PROFILESCAN
17	101	T59 S67		M1-S40; M1-S28	MOTIFS SPSCAN HMM

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Peptides and Other Signature Sequences	Potential Signal Peptides and Other Sequences	Analytical Methods
18	540	S42 S38 T80 T172 S179 S326 S519 S531 T131 T250 T278 Y99	N78 N88 N170 N347 N448 N457	M1-C29; M1-I31 Transmembrane domains: V301-N327 L411-I428 A140-L163 D366-I385 Y99-Y122 F496-C513 I56-Y73		MOTIFS SPSCAN HMM PRINTS
19	108	S34 S64 S29 S47 Y96		M1-R22; M1-S23		MOTIFS SPSCAN HMM
20	114	S108 S5 S73 T85		M1-T32		MOTIFS SPSCAN HMM
21	114	T8	N6	M1-A34		MOTIFS SPSCAN HMM
22	287	S25 T75 T134 T139 S225 T255 S254		M1-P20; M1-A22 RGD tripeptide: R172-D174		MOTIFS SPSCAN HMM

Table 3

Nucleotide SEQ ID NO:	Selected Fragment (s)	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
23	110-154 542-586	Musculoskeletal (0.750) Reproductive (0.250)	Inflammation (0.750) Cancer (0.250)	PBLUESCRIPT
24	230-271	Reproductive (0.333) Musculoskeletal (0.190) Cardiovascular (0.143)	Cancer (0.476) Inflammation (0.333)	PBLUESCRIPT
25	109-153 649-693	Reproductive (0.314) Nervous (0.235) Gastrointestinal (0.157)	Cancer (0.608) Inflammation (0.196) Cell Proliferation (0.118)	PSPORT1
26	116-160	Reproductive (0.333) Nervous (0.178) Cardiovascular (0.156)	Cancer (0.600) Cell Proliferation (0.178) Inflammation (0.133)	PSPORT1
27	228-272	Reproductive (0.333) Cardiovascular (0.244) Gastrointestinal (0.111)	Cancer (0.667) Cell Proliferation (0.133) Inflammation (0.089)	PSPORT1
28	1945-1989	Nervous (0.301) Reproductive (0.219) Gastrointestinal (0.137)	Cancer (0.452) Inflammation (0.205) Trauma (0.164)	PINCY
29	271-315	Gastrointestinal (0.800) Nervous (0.200)	Trauma (0.600) Cancer (0.200) Inflammation (0.200)	PINCY
30	218-262	Reproductive (0.249) Nervous (0.195) Gastrointestinal (0.136)	Cancer (0.521) Inflammation (0.207) Cell Proliferation (0.172)	PINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
31	1190-1234	Nervous (0.324) Reproductive (0.250) Hematopoietic/Immune (0.118)	Cancer (0.426) Inflammation (0.279) Cell Proliferation (0.147)	PSPORT1
32	487-531	Gastrointestinal (0.333) Nervous (0.259) Reproductive (0.111)	Cancer (0.593) (0.148) Inflammation Cell Proliferation (0.111)	PINCY
33	1513-1557	Reproductive (0.255) Nervous (0.216) Hematopoietic/Immune (0.176)	Cancer (0.373) Inflammation (0.255) Cell Proliferation (0.196)	PINCY
34	270-317	Hematopoietic/Immune (0.455) Musculoskeletal (0.182) Nervous (0.182)	Inflammation (0.545) Cancer (0.273) Cell Proliferation (0.091)	PINCY
35	1299-1343 1956-2000	Musculoskeletal (0.519) Nervous (0.148) Cardiovascular (0.111)	Inflammation (0.481) Cancer (0.296) Trauma (0.074)	PINCY
36	651-695	Nervous (0.250) Developmental (0.167) Gastrointestinal (0.167)	Cell Proliferation (0.667) Inflammation (0.250) Cancer (0.167)	PINCY
37	218-262	Nervous (0.818) Gastrointestinal (0.091) Reproductive (0.091)	Cancer (0.545) Cell Proliferation (0.091) Inflammation (0.091)	PINCY
38	290-334	Hematopoietic/Immune (0.250) Urologic (0.250) Developmental (0.125)	Cell Proliferation (0.375) Inflammation (0.250) Cancer (0.125)	PINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
39	434-478	Reproductive (1.00)	Cancer (1.00)	pINCY
40	326-370	Nervous (0.242) Reproductive (0.220) Gastrointestinal (0.121)	Cancer (0.462) Inflammation (0.280) Cell Proliferation (0.121)	pINCY
41	165-209	Nervous (0.333) Gastrointestinal (0.200) Cardiovascular (0.133)	Cancer (0.467) Cell Proliferation (0.200) Inflammation (0.133)	pINCY
42	273-317	Hematopoietic/Immune (0.312) Nervous (0.188) Reproductive (0.167)	Cancer (0.354) Inflammation (0.312) Cell Proliferation (0.146)	pINCY
43	273-317	Cardiovascular (1.000)	Inflammation (0.500)	pINCY
44	435-479	Reproductive (0.297) Nervous (0.217) Gastrointestinal (0.109)	Cancer (0.464) Cell Proliferation (0.196) Inflammation (0.167)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
23	SYNORAB01	This library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.
24	LUNGNOT02	This library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male who died of a subarachnoid hemorrhage.
25	PROSNOT02	This library was constructed using RNA isolated from diseased prostate tissue removed from a 50-year-old Caucasian male during a retropubic prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma Gleason grade 3+3. Patient history included dysuria, carcinoma in situ of prostate, coronary atherosclerosis, and hyperlipidemia.
26	BRAITUT03	This library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
27	BRSTNOT02	This library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated proliferative fibrocystic changes characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Pathology for the associated tumor tissue indicated an invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign neoplasm. Family history included cardiovascular and cerebrovascular disease.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
28	PROSTUT08	This library was constructed using RNA isolated from prostate tumor tissue removed from a 60-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Patient history included a kidney cyst, and hematuria. Family history included tuberculosis, cerebrovascular disease, and arteriosclerotic coronary artery disease.
29	DUODNOT02	This library was constructed using RNA isolated from duodenal tissue of an 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
30	BRSTTUT08	This library was constructed using RNA isolated from breast tumor tissue removed from a 45-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive nuclear grade 2-3 adenocarcinoma, ductal type, with 3 of 23 lymph nodes positive for metastatic disease. Greater than 50% of the tumor volume was <i>in situ</i> , both comedo and non-comedo types. Immunostains were positive for estrogen/progesterone receptors, and uninvolved tissue showed proliferative changes. The patient concurrently underwent a total abdominal hysterectomy. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
31	PITUNOT03	This library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male who died from colon cancer.
32	PANCTUT02	This library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
33	ENDCMNOT03	This library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
34	GBLANOT01	This library was constructed using RNA isolated from diseased gallbladder tissue removed from a 53-year-old Caucasian female during a cholecystectomy. Pathology indicated mild chronic cholecystitis and cholelithiasis with approximately 150 mixed gallstones. Family history included benign hypertension.
35	KIDNFETO2	This library was constructed using RNA isolated from kidney tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
36	TLXNOT05	This library was constructed using RNA isolated from nonactivated Th2 cells. The cells were differentiated from umbilical cord CD4 T cells with IL-4 in the presence of anti-IL-12 antibodies and B7-transfected COS cells.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
37	BRAINOT20	This library was constructed using RNA isolated from diseased brain tissue removed from the left temporal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology for the left temporal lobe, including the mesial temporal structures, indicated focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. The left frontal lobe showed a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This frontal lobe tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. GFAP was positive for astrocytes. Family history included brain cancer.
38	TESTNOT07	This library was constructed using RNA isolated from testicular tissue removed from a 31-year-old Caucasian male during a unilateral orchectomy (excision of testis). Pathology indicated a mass containing a large subcapsular hematoma with laceration of the tunica albuginea. The surrounding testicular parenchyma was extensively necrotic. The patient presented with a trunk injury.
39	PROSTUS08	This subtracted library was constructed using 2.36 million clones from a prostate tumor library and was subjected to one round of subtractive hybridization with 48,000 clones from a prostate library. The starting library for subtraction was constructed using RNA isolated from a prostate tumor removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenocarcinoma (Gleason grade 3+3) and adenofibromatous hyperplasia. The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestos, and thrombophlebitis. Family history included multiple myeloma, hyperlipidemia, and rheumatoid arthritis. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954 and Bonaldo, et al. Genome Research (1996) 6:791.
40	BRAIDIT01	This library was constructed using RNA isolated from diseased brain tissue. Patient history included multiple sclerosis, type II lesion.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
41	293TF5T01	This library was constructed using RNA isolated from a transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue transfected with bgal. The cells were transformed with adenovirus 5 DNA.
42	KIDNTUT16	This library was constructed using RNA isolated from left pole kidney tumor tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Pathology indicated grade 2 renal cell carcinoma. Patient history included hyperlipidemia, cardiac dysrhythmia, menorrhagia, cerebrovascular disease, atherosclerotic coronary artery disease, and tobacco abuse. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
43	HEAANOT01	This library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Previous surgeries included cardiac catheterization. Family history included atherosclerotic coronary artery disease.
44	THYMMNOT08	This library was constructed using RNA isolated from thymus tissue removed from a 4-month-old Caucasian male during a total thymectomy and open heart repair of atrioventricular canal defect using hypothermia. Patient presented with a congenital heart anomaly, congestive heart failure, and Down syndrome. Patient history included abnormal thyroid function study, premature birth, and right and left heart angiography.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	<i>ESTs</i> : Probability value= 1.0E-8 or less <i>Assembled ESTs</i> : Probability value= 1.0E-10 or less <i>Full Length sequences</i> : Probability value= 1.0E-6
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	<i>ESTs</i> : fasta E value= 1.06E-6 <i>Assembled ESTs</i> : fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value= 1.0E-8 or less <i>Full Length sequences</i> : fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score $\geq$ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audit (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motif	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <u>supra</u> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
  - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22,
  - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22,
  - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, and
  - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-22.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID NO:23-44.
- 20 5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
6. A cell transformed with a recombinant polynucleotide of claim 5.
- 25 7. A transgenic organism comprising a recombinant polynucleotide of claim 5.
8. A method for producing a polypeptide of claim 1, the method comprising:
  - 30 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
  - b) recovering the polypeptide so expressed.
- 35 9. An isolated antibody which specifically binds to a polypeptide of claim 1.

10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44,
- b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a 5 polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44,
- c) a polynucleotide sequence complementary to a),
- d) a polynucleotide sequence complementary to b), and
- e) an RNA equivalent of a)-d).

10 11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 10.

12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:

- 15 a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if 20 present, the amount thereof.

13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.

14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.

25 15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

30 16. A method for treating a disease or condition associated with decreased expression of functional HSECP, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 15.

35 17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

5

19. A method for treating a disease or condition associated with decreased expression of functional HSECP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

10 20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

15 21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with overexpression of functional HSECP, comprising administering to a patient in need of such treatment a pharmaceutical 20 composition of claim 21.

25 23. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

## SEQUENCE LISTING

&lt;110&gt; INCYTE PHARMACEUTICALS, INC.

TANG, Y. Tom  
LAL, Preeti  
BAUGHN, Mariah R.  
YUE, Henry  
AU-YOUNG, Janice  
LU, Dyung Aina M.  
AZIMZAI, Yalda

&lt;120&gt; HUMAN SECRETORY PROTEINS

&lt;130&gt; PF-0675 PCT

<140> To Be Assigned  
<141> Herewith

<150> 60/123,117  
<151> 1999-03-05

&lt;160&gt; 44

&lt;170&gt; PERL Program

<210> 1  
<211> 182  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 078811CD1

<400> 1  
Met Arg Ser Thr Ile Leu Leu Phe Cys Leu Leu Gly Ser Thr Arg  
1 5 10 15  
Ser Leu Pro Val Phe Pro Ser Leu Ser Leu Ile Pro Leu Thr Gln  
20 25 30  
Met Leu Thr Leu Gly Pro Asp Leu His Leu Leu Asn Pro Ala Ala  
35 40 45  
Gly Met Thr Pro Gly Thr Gln Thr His Pro Leu Thr Leu Gly Gly  
50 55 60  
Leu Asn Val Gln Gln Leu His Pro His Val Leu Pro Ile Phe  
65 70 75  
Val Thr Gln Leu Gly Ala Pro Gly His Tyr Pro Lys Leu Arg Gly  
80 85 90  
Ile Ala Thr Asn Leu His Glu Pro His His Pro Phe Leu Val Pro  
95 100 105  
Arg Glu Ala Ser Leu Pro Thr Ser Gln Ala Gly Ala Asn Pro Asp  
110 115 120  
Val Gln Asp Gly Ser Leu Pro Ala Gly Gly Ala Gly Val Asn Pro  
125 130 135  
Ala Thr Gln Gly Thr Pro Ala Gly Arg Leu Pro Thr Pro Ser Gly  
140 145 150  
Thr Asp Asp Asp Phe Ala Val Thr Thr Pro Ala Gly Ile Gln Arg  
155 160 165  
Ser Thr His Ala Ile Glu Glu Ala Thr Thr Glu Ser Ala Asn Gly  
170 175 180  
Ile Gln

<210> 2  
<211> 125

<212> PRT  
 <213> Homo sapiens

<220> -  
 <221> misc\_feature  
 <223> Incyte ID No: 371156CD1

<400> 2  
 Met Val Cys Glu Asp Ala Pro Ser Phe Gln Met Ala Trp Glu Ser  
 1 5 10 15  
 Gln Met Ala Trp Glu Arg Gly Pro Ala Leu Leu Cys Cys Val Leu  
 20 25 30  
 Ser Ala Ser Gln Leu Ser Ser Gln Asp Gln Asp Pro Leu Gly His  
 35 40 45  
 Ile Lys Ser Leu Leu Tyr Pro Phe Gly Phe Pro Val Glu Leu Pro  
 50 55 60  
 Arg Pro Gly Pro Thr Gly Ala Tyr Lys Lys Val Lys Asn Glu Asp  
 65 70 75  
 Gln Thr Thr Ser Ser Glu Leu Leu Arg Lys Gln Thr Ser His Phe  
 80 85 90  
 Asn Gln Arg Gly His Arg Ala Arg Ser Lys Leu Leu Ala Ser Arg  
 95 100 105  
 Gln Ile Pro Asp Arg Thr Phe Lys Cys Gly Lys Trp Leu Pro Gln  
 110 115 120  
 Val Pro Ser Pro Val  
 125

<210> 3  
 <211> 320  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 584050CD1

<400> 3  
 Met Ala Gly Leu Ala Ala Arg Leu Val Leu Leu Ala Gly Ala Ala  
 1 5 10 15  
 Ala Leu Ala Ser Gly Ser Gln Gly Asp Arg Glu Pro Val Tyr Arg  
 20 25 30  
 Asp Cys Val Leu Gln Cys Glu Glu Gln Asn Cys Ser Gly Gly Ala  
 35 40 45  
 Leu Asn His Phe Arg Ser Arg Gln Pro Ile Tyr Met Ser Leu Ala  
 50 55 60  
 Gly Trp Thr Cys Arg Asp Asp Cys Lys Tyr Glu Cys Met Trp Val  
 65 70 75  
 Thr Val Gly Leu Tyr Leu Gln Glu Gly His Lys Val Pro Gln Phe  
 80 85 90  
 His Gly Lys Trp Pro Phe Ser Arg Phe Leu Phe Phe Gln Glu Pro  
 95 100 105  
 Ala Ser Ala Val Ala Ser Phe Leu Asn Gly Leu Ala Ser Leu Val  
 110 115 120  
 Met Leu Cys Arg Tyr Arg Thr Phe Val Pro Ala Ser Ser Pro Met  
 125 130 135  
 Tyr His Thr Cys Val Ala Phe Ala Trp Val Ser Leu Asn Ala Trp  
 140 145 150  
 Phe Trp Ser Thr Val Phe His Thr Arg Asp Thr Asp Leu Thr Glu  
 155 160 165  
 Lys Met Asp Tyr Phe Cys Ala Ser Thr Val Ile Leu His Ser Ile  
 170 175 180  
 Tyr Leu Cys Cys Val Arg Thr Val Gly Leu Gln His Pro Ala Val  
 185 190 195  
 Val Ser Ala Phe Arg Ala Leu Leu Leu Met Leu Thr Val His

200	205	210
Val Ser Tyr Leu Ser	Leu Ile Arg Phe Asp Tyr Gly Tyr Asn	Leu
215	220	225
Val Ala Asn Val Ala	Ile Gly Leu Val Asn Val Val Trp Trp	Leu
230	235	240
Ala Trp Cys Leu Trp	Asn Gln Arg Arg Leu Pro His Val Arg	Lys
245	250	255
Cys Val Val Val Val	Leu Leu Leu Gln Gly Leu Ser Leu Leu	Glu
260	265	270
Leu Leu Asp Phe Pro	Pro Leu Phe Trp Val Leu Asp Ala His	Ala
275	280	285
Ile Trp His Ile Ser	Thr Ile Pro Val His Val Leu Phe Phe	Ser
290	295	300
Phe Leu Glu Asp Asp	Ser Leu Tyr Leu Leu Lys Glu Ser Glu Asp	
305	310	315
Lys Phe Lys Leu Asp		
320		

<210> 4  
<211> 234  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 863808CD1

<400> 4  
Met Gly Pro Gly Gly Arg Val Ala Arg Leu Leu Ala Pro Leu Met  
1 5 10 15  
Trp Arg Arg Ala Val Ser Ser Val Ala Gly Ser Ala Val Gly Ala  
20 25 30  
Glu Pro Gly Leu Arg Leu Leu Ala Val Gln Arg Leu Pro Val Gly  
35 40 45  
Ala Ala Phe Cys Arg Ala Cys Gln Thr Pro Asn Phe Val Arg Gly  
50 55 60  
Leu His Ser Glu Pro Gly Leu Glu Glu Arg Ala Glu Gly Thr Val  
65 70 75  
Asn Glu Gly Arg Pro Glu Ser Asp Ala Ala Asp His Thr Gly Pro  
80 85 90  
Lys Phe Asp Ile Asp Met Met Val Ser Leu Leu Arg Gln Glu Asn  
95 100 105  
Ala Arg Asp Ile Cys Val Ile Gln Val Pro Pro Glu Met Arg Tyr  
110 115 120  
Thr Asp Tyr Phe Val Ile Val Ser Gly Thr Ser Thr Arg His Leu  
125 130 135  
His Ala Met Ala Phe Tyr Val Val Lys Met Tyr Lys His Leu Lys  
140 145 150  
Cys Lys Arg Asp Pro His Val Lys Ile Glu Gly Lys Asp Thr Asp  
155 160 165  
Asp Trp Leu Cys Val Asp Phe Gly Ser Met Val Ile His Leu Met  
170 175 180  
Leu Pro Glu Thr Arg Glu Ile Tyr Glu Leu Glu Lys Leu Trp Thr  
185 190 195  
Leu Arg Ser Tyr Asp Asp Gln Leu Ala Gln Ile Ala Pro Glu Thr  
200 205 210  
Val Pro Glu Asp Phe Ile Leu Gly Ile Glu Asp Asp Thr Ser Ser  
215 220 225  
Val Thr Pro Val Glu Leu Lys Cys Glu  
230

<210> 5  
 <211> 278  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 978433CD1

<400> 5  
 Met Gln Pro Ala Ala Ala Ser Glu Arg Gly Gly Ala Asp Ala Asp  
 1 5 10 15  
 His Val Pro Leu Leu Gly Leu Leu Arg Leu Gln Leu Arg Ala Ala  
 20 25 30  
 Arg Gln Pro Gly Ala Met Arg Pro Gln Gly Pro Ala Ala Ser Pro  
 35 40 45  
 Gln Arg Leu Arg Gly Leu Leu Leu Leu Leu Leu Gln Leu Pro  
 50 55 60  
 Ala Pro Ser Ser Ala Ser Glu Ile Pro Lys Gly Lys Gln Lys Ala  
 65 70 75  
 Gln Leu Arg Gln Arg Glu Val Val Asp Leu Tyr Asn Gly Met Cys  
 80 85 90  
 Leu Gln Gly Pro Ala Gly Val Pro Gly Arg Asp Gly Ser Pro Gly  
 95 100 105  
 Ala Asn Gly Ile Pro Gly Thr Pro Gly Ile Pro Gly Arg Asp Gly  
 110 115 120  
 Phe Lys Gly Glu Lys Gly Glu Cys Leu Arg Glu Ser Phe Glu Glu  
 125 130 135  
 Ser Trp Thr Pro Asn Tyr Lys Gln Cys Ser Trp Ser Ser Leu Asn  
 140 145 150  
 Tyr Gly Ile Asp Leu Gly Lys Ile Ala Glu Cys Thr Phe Thr Lys  
 155 160 165  
 Met Arg Ser Asn Ser Ala Leu Arg Val Leu Phe Ser Gly Ser Leu  
 170 175 180  
 Arg Leu Lys Cys Arg Asn Ala Cys Cys Gln Arg Trp Tyr Phe Thr  
 185 190 195  
 Phe Asn Gly Ala Glu Cys Ser Gly Pro Leu Pro Ile Glu Ala Ile  
 200 205 210  
 Ile Tyr Leu Asp Gln Gly Ser Pro Glu Met Asn Ser Thr Ile Asn  
 215 220 225  
 Ile His Arg Thr Ser Ser Val Glu Gly Leu Cys Glu Gly Ile Gly  
 230 235 240  
 Ala Gly Leu Val Asp Val Ala Ile Trp Val Gly Thr Cys Ser Asp  
 245 250 255  
 Tyr Pro Lys Gly Asp Ala Ser Thr Gly Trp Asn Ser Val Ser Arg  
 260 265 270  
 Ile Ile Ile Glu Glu Leu Pro Lys  
 275

<210> 6  
 <211> 136  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1655369CD1

<400> 6  
 Met Pro Pro Gly Gly Leu Gly Ala Cys Ala Val Thr Pro Ala Pro  
 1 5 10 15  
 Gly Glu Glu Arg Thr Gln Pro Gly Glu Leu Gly Gln Gly Leu His  
 20 25 30  
 Met Ala Gln Gly Gln Gln Met Leu Ala Gly Gln Leu Leu Pro Met

35	40	45
Leu Thr Leu Leu Pro Pro Ser Phe Pro	Leu Pro His Pro Thr	Leu
50	55	60
Gly Pro Arg Arg His Ala Ser Leu Thr	Gln Leu Gly Pro Ala Phe	
65	70	75
Trp Met Ala Trp Gly Arg Pro Trp Ala	His Leu Gly Pro Gly Gln	
80	85	90
Pro Leu Gly Gln Leu Trp Lys Ser Ser	Val Glu Glu His Leu Leu	
95	100	105
Ala Ala Trp Leu Gln Pro Leu Ala Leu	Leu Glu Trp Ser Leu Gly	
110	115	120
Ala Ser Ala Leu Ser Ala Leu Gly Thr	Ser His Pro Leu Gly Leu	
125	130	135
Gln		

<210> 7  
 <211> 109  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1703244CD1

<400> 7

Met Leu Met Tyr Met	Phe Tyr Val Leu Pro	Phe Cys Gly Leu Ala	
1	5	10	15
Ala Tyr Ala Leu Thr	Phe Pro Gly Cys Ser	Trp Leu Pro Asp Trp	
20	25	30	
Ala Leu Val Phe Ala	Gly Gly Ile Gly Gln	Ala Gln Phe Ser His	
35	40	45	
Met Gly Ala Ser Met	His Leu Arg Thr Pro	Phe Thr Tyr Arg Val	
50	55	60	
Pro Glu Asp Thr Trp	Gly Cys Phe Phe Val	Cys Asn Leu Leu Tyr	
65	70	75	
Ala Leu Gly Pro His	Leu Leu Ala Tyr Arg	Cys Leu Gln Trp Pro	
80	85	90	
Ala Phe Phe His Gln	Pro Pro Pro Ser Asp	Pro Leu Ala Leu His	
95	100	105	
Lys Lys Gln His			

<210> 8  
 <211> 262  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1730819CD1

<400> 8

Met Ala Ala Ala Ser Ala Gly Ala Thr Arg	Leu Leu Leu Leu		
1	5	10	15
Leu Met Ala Val Ala Ala Pro Ser Arg Ala	Arg Gly Ser Gly Cys		
20	25	30	
Arg Ala Gly Thr Gly Ala Arg Gly Ala	Gly Ala Glu Gly Arg Glu		
35	40	45	
Gly Glu Ala Cys Gly Thr Val Gly Leu Leu	Leu Glu His Ser Phe		
50	55	60	
Glu Ile Asp Asp Ser Ala Asn Phe Arg	Lys Arg Gly Ser Leu Leu		
65	70	75	

Trp	Asn	Gln	Gln	Asp	Gly	Thr	Leu	Ser	Leu	Ser	Gln	Arg	Gln	Leu
80							85							90
Ser	Glu	Glu	Glu	Arg	Gly	Arg	Leu	Arg	Asp	Val	Ala	Ala	Leu	Asn
95							100							105
Gly	Leu	Tyr	Arg	Val	Arg	Ile	Pro	Arg	Arg	Pro	Gly	Ala	Leu	Asp
110							115							120
Gly	Leu	Glu	Ala	Gly	Gly	Tyr	Val	Ser	Ser	Phe	Val	Pro	Ala	Cys
125							130							135
Ser	Leu	Val	Glu	Ser	His	Leu	Ser	Asp	Gln	Leu	Thr	Leu	His	Val
140							145							150
Asp	Val	Ala	Gly	Asn	Val	Val	Gly	Val	Ser	Val	Val	Thr	His	Pro
155							160							165
Gly	Gly	Cys	Arg	Gly	His	Glu	Val	Glu	Asp	Val	Asp	Leu	Glu	Leu
170							175							180
Phe	Asn	Thr	Ser	Val	Gln	Leu	Gln	Pro	Pro	Thr	Thr	Ala	Pro	Gly
185							190							195
Pro	Glu	Thr	Ala	Ala	Phe	Ile	Glu	Arg	Leu	Glu	Met	Glu	Gln	Ala
200							205							210
Gln	Lys	Ala	Lys	Asn	Pro	Gln	Glu	Gln	Lys	Ser	Phe	Phe	Ala	Lys
215							220							225
Tyr	Trp	Met	Tyr	Ile	Ile	Pro	Val	Val	Leu	Phe	Leu	Met	Met	Ser
230							235							240
Gly	Ala	Pro	Asp	Thr	Gly	Gly	Gln	Gly	Gly	Gly	Gly	Cys	Gly	
245							250							255
Gly	Gly	Gly	Gly	Ser	Gly	Arg								
260														

<210> 9  
 <211> 384  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1757161CD1

<400> 9															
Met	Ala	Glu	Gln	Thr	Tyr	Ser	Trp	Ala	Tyr	Ser	Leu	Val	Asp	Ser	
1								5						15	
Ser	Gln	Val	Val	Ser	Thr	Phe	Leu	Ile	Ser	Ile	Leu	Leu	Ile	Val	Tyr
								20							30
Gly	Ser	Phe	Arg	Ser	Leu	Asn	Met	Asp	Phe	Glu	Asn	Gln	Asp	Lys	
								35							45
Glu	Lys	Asp	Ser	Asn	Ser	Ser	Ser	Gly	Ser	Phe	Asn	Gly	Asn	Ser	
								50							60
Thr	Asn	Asn	Ser	Ile	Gln	Thr	Ile	Asp	Ser	Thr	Gln	Ala	Leu	Phe	
								65							75
Leu	Pro	Ile	Gly	Ala	Ser	Val	Ser	Leu	Leu	Val	Met	Phe	Phe		
								80							90
Phe	Asp	Ser	Val	Gln	Val	Val	Phe	Thr	Ile	Cys	Thr	Ala	Val	Leu	
								95							105
Ala	Thr	Ile	Ala	Phe	Ala	Phe	Leu	Leu	Leu	Pro	Met	Cys	Gln	Tyr	
								110							120
Leu	Thr	Arg	Pro	Cys	Ser	Pro	Gln	Asn	Lys	Ile	Ser	Phe	Gly	Cys	
								125							135
Cys	Gly	Arg	Phe	Thr	Ala	Ala	Glu	Leu	Leu	Ser	Phe	Ser	Leu	Ser	
								140							150
Val	Met	Leu	Val	Leu	Ile	Trp	Val	Leu	Thr	Gly	His	Trp	Leu	Leu	
								155							165
Met	Asp	Ala	Leu	Ala	Met	Gly	Leu	Cys	Val	Ala	Met	Ile	Ala	Phe	
								170							180
Val	Arg	Leu	Pro	Ser	Leu	Lys	Val	Ser	Cys	Leu	Leu	Leu	Ser	Gly	
								185							195
Leu	Leu	Ile	Tyr	Asp	Val	Phe	Trp	Val	Phe	Phe	Ser	Ala	Tyr	Ile	

	200	205	210
Phe Asn Ser Asn	Val Met Val Lys Val	Ala Thr Gln Pro Ala Asp	
	215	220	225
Asn Pro Leu Asp	Val Leu Ser Arg Lys	Leu His Leu Gly Pro Asn	
	230	235	240
Val Gly Arg Asp	Val Pro Arg Leu Ser	Leu Pro Gly Lys Leu Val	
	245	250	255
Phe Pro Ser Ser	Thr Gly Ser His Phe Ser	Met Leu Gly Ile Gly	
	260	265	270
Asp Ile Val Met	Pro Gly Leu Leu Leu	Cys Phe Val Leu Arg Tyr	
	275	280	285
Asp Asn Tyr Lys	Lys Gln Ala Ser Gly	Asp Ser Cys Gly Ala Pro	
	290	295	300
Gly Pro Ala Asn	Ile Ser Gly Arg Met	Gln Lys Val Ser Tyr Phe	
	305	310	315
His Cys Thr Leu	Ile Gly Tyr Phe Val	Gly Leu Leu Thr Ala Thr	
	320	325	330
Val Ala Ser Arg	Ile His Arg Ala Ala	Gln Pro Ala Leu Leu Tyr	
	335	340	345
Leu Val Pro Phe	Thr Leu Leu Pro Leu	Leu Thr Met Ala Tyr Leu	
	350	355	360
Lys Gly Asp Leu	Arg Arg Met Trp Ser	Glu Pro Phe His Ser Lys	
	365	370	375
Ser Ser Ser Ser	Arg Phe Leu Glu Val		
	380		

<210> 10  
<211> 244  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1976095CD1

	<400> 10		
Met Asp Ile Leu Val	Pro Leu Leu Gln Leu Leu Val Leu Leu Leu		
1	5	10	15
Thr Leu Pro Leu His	Leu Met Ala Leu Leu Gly Cys Trp Gln Pro		
	20	25	30
Leu Cys Lys Ser Tyr	Phe Pro Tyr Leu Met Ala Val Leu Thr Pro		
	35	40	45
Lys Ser Asn Arg Lys	Met Glu Ser Lys Lys Arg Glu Leu Phe Ser		
	50	55	60
Gln Ile Lys Gly Leu	Thr Gly Ala Ser Gly Lys Val Ala Leu Leu		
	65	70	75
Glu Leu Gly Cys Gly	Thr Gly Ala Asn Phe Gln Phe Tyr Pro Pro		
	80	85	90
Gly Cys Arg Val	Thr Cys Leu Asp Pro Asn Pro His Phe Glu Lys		
	95	100	105
Phe Leu Thr Lys	Ser Met Ala Glu Asn Arg His Leu Gln Tyr Glu		
	110	115	120
Arg Phe Val Val	Ala Pro Gly Glu Asp Met Arg Gln Leu Ala Asp		
	125	130	135
Gly Ser Met Asp	Val Val Val Cys Thr Leu Val Leu Cys Ser Val		
	140	145	150
Gln Ser Pro Arg	Lys Val Leu Gln Glu Val Arg Arg Val Leu Arg		
	155	160	165
Pro Gly Gly Val	Leu Phe Phe Trp Glu His Val Ala Glu Pro Tyr		
	170	175	180
Gly Ser Trp Ala	Phe Met Trp Gln Gln Val Phe Glu Pro Thr Trp		
	185	190	195
Lys His Ile Gly	Asp Gly Cys Cys Leu Thr Arg Glu Thr Trp Lys		
	200	205	210

Asp Leu Glu Asn Ala Gln Phe Ser Glu Ile Gln Met Glu Arg Gln  
 215 220 225  
 Pro Pro Pro Leu Lys Trp Leu Pro Val Gly Pro His Ile Met Gly  
 230 235 240  
 Lys Ala Val Lys

<210> 11  
 <211> 326  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2169991CD1

<400> 11  
 Met Arg Thr Glu Ala Gln Val Pro Ala Leu Gln Pro Pro Glu Pro  
 1 5 10 15  
 Gly Leu Glu Gly Ala Met Gly His Arg Thr Leu Val Leu Pro Trp  
 20 25 30  
 Val Leu Leu Thr Leu Cys Val Thr Ala Gly Thr Pro Glu Val Trp  
 35 40 45  
 Val Gln Val Arg Met Glu Ala Thr Glu Leu Ser Ser Phe Thr Ile  
 50 55 60  
 Arg Cys Gly Phe Leu Gly Ser Gly Ser Ile Ser Leu Val Thr Val  
 65 70 75  
 Ser Trp Gly Gly Pro Asn Gly Ala Gly Gly Thr Thr Leu Ala Val  
 80 85 90  
 Leu His Pro Glu Arg Gly Ile Arg Gln Trp Ala Pro Ala Arg Gln  
 95 100 105  
 Ala Arg Trp Glu Thr Gln Ser Ser Ile Ser Leu Ile Leu Glu Gly  
 110 115 120  
 Ser Gly Ala Ser Ser Pro Cys Ala Asn Thr Thr Phe Cys Cys Lys  
 125 130 135  
 Phe Ala Ser Phe Pro Glu Gly Ser Trp Glu Ala Cys Gly Ser Leu  
 140 145 150  
 Pro Pro Ser Ser Asp Pro Gly Leu Ser Ala Pro Pro Thr Pro Ala  
 155 160 165  
 Pro Ile Leu Arg Ala Asp Leu Ala Gly Ile Leu Gly Val Ser Gly  
 170 175 180  
 Val Leu Leu Phe Gly Cys Val Tyr Leu Leu His Leu Leu Arg Arg  
 185 190 195  
 His Lys His Arg Pro Ala Pro Arg Leu Gln Pro Ser Arg Thr Ser  
 200 205 210  
 Pro Gln Ala Pro Arg Ala Arg Ala Trp Ala Pro Ser Gln Ala Ser  
 215 220 225  
 Gln Ala Ala Leu His Val Pro Tyr Ala Thr Ile Asn Thr Ser Cys  
 230 235 240  
 Arg Pro Ala Thr Leu Asp Thr Ala His Pro His Gly Gly Pro Ser  
 245 250 255  
 Trp Trp Ala Ser Leu Pro Thr His Ala Ala His Arg Pro Gln Gly  
 260 265 270  
 Pro Ala Ala Trp Ala Ser Thr Pro Ile Pro Ala Arg Gly Ser Phe  
 275 280 285  
 Val Ser Val Glu Asn Gly Leu Tyr Ala Gln Ala Gly Glu Arg Pro  
 290 295 300  
 Pro His Thr Gly Pro Gly Leu Thr Leu Phe Pro Asp Pro Arg Gly  
 305 310 315  
 Pro Arg Ala Met Glu Gly Pro Leu Gly Val Arg  
 320 325

<210> 12  
 <211> 105  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2616827CD1

<400> 12

Met	Asn	Leu	Gly	Val	Ser	Met	Leu	Arg	Ile	Leu	Phe	Leu	Leu	Asp
1						5			10					15
Val	Gly	Gly	Ala	Gln	Val	Leu	Ala	Thr	Gly	Lys	Thr	Pro	Gly	Ala
						20			25					30
Glu	Ile	Asp	Phe	Lys	Tyr	Ala	Leu	Ile	Gly	Thr	Ala	Val	Gly	Val
						35			40					45
Ala	Ile	Ser	Ala	Gly	Phe	Leu	Ala	Leu	Lys	Ile	Cys	Met	Ile	Arg
						50			55					60
Arg	His	Leu	Phe	Asp	Asp	Asp	Ser	Ser	Asp	Leu	Lys	Ser	Thr	Pro
						65			70					75
Gly	Gly	Leu	Ser	Asp	Thr	Ile	Pro	Leu	Lys	Lys	Arg	Ala	Pro	Arg
						80			85					90
Arg	Asn	His	Asn	Phe	Ser	Lys	Arg	Asp	Ala	Gln	Val	Ile	Glu	Leu
						95			100					105

<210> 13  
 <211> 626  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2991370CD1

<400> 13

Met	Ala	Pro	Ser	Ala	Asp	Pro	Gly	Met	Ser	Arg	Met	Leu	Pro	Phe
1						5			10					15
Leu	Leu	Leu	Leu	Trp	Phe	Leu	Pro	Ile	Thr	Glu	Gly	Ser	Gln	Arg
						20			25					30
Ala	Glu	Pro	Met	Phe	Thr	Ala	Val	Thr	Asn	Ser	Val	Leu	Pro	Pro
						35			40					45
Asp	Tyr	Asp	Ser	Asn	Pro	Thr	Gln	Leu	Asn	Tyr	Gly	Val	Ala	Val
						50			55					60
Thr	Asp	Val	Asp	His	Asp	Gly	Asp	Phe	Glu	Ile	Val	Val	Ala	Gly
						65			70					75
Tyr	Asn	Gly	Pro	Asn	Leu	Val	Leu	Lys	Tyr	Asp	Arg	Ala	Gln	Lys
						80			85					90
Arg	Leu	Val	Asn	Ile	Ala	Val	Asp	Glu	Arg	Ser	Ser	Pro	Tyr	Tyr
						95			100					105
Ala	Leu	Arg	Asp	Arg	Gln	Gly	Asn	Ala	Ile	Gly	Val	Thr	Ala	Cys
						110			115					120
Asp	Ile	Asp	Gly	Asp	Gly	Arg	Glu	Glu	Ile	Tyr	Phe	Leu	Asn	Thr
						125			130					135
Asn	Asn	Ala	Phe	Ser	Gly	Val	Ala	Thr	Tyr	Thr	Asp	Lys	Leu	Phe
						140			145					150
Lys	Phe	Arg	Asn	Asn	Arg	Trp	Glu	Asp	Ile	Leu	Ser	Asp	Glu	Val
						155			160					165
Asn	Val	Ala	Arg	Gly	Val	Ala	Ser	Leu	Phe	Ala	Gly	Arg	Ser	Val
						170			175					180
Ala	Cys	Val	Asp	Arg	Lys	Gly	Ser	Gly	Arg	Tyr	Ser	Ile	Tyr	Ile
						185			190					195
Ala	Asn	Tyr	Ala	Tyr	Gly	Asn	Val	Gly	Pro	Asp	Ala	Leu	Ile	Glu
						200			205					210
Met	Asp	Pro	Glu	Ala	Ser	Asp	Leu	Ser	Arg	Gly	Ile	Leu	Ala	Leu

215	Arg Asp Val Ala Ala	Glu Ala Gly Val	Ser Lys Tyr Thr Gly	220 225	
230	Arg Gly Val Ser Val	Gly Pro Ile Leu	Ser Ser Ser Ala Ser	235 240	
245	Ile Phe Cys Asp Asn	Glu Asn Gly Pro	Asn Phe Leu Phe His	250 255	
260	Arg Gly Asp Gly Thr	Phe Val Asp Ala	Ala Ala Ser Ala Gly	265 270	
275	Asp Asp Pro His Gln	Val His Gly Arg Gly	Val Ala Leu Ala Asp	280 285	
290	Asn Arg Asp Gly Lys	Tyr Val Asp Ile Val	Phe 295	300	
305	Pro His Arg Leu Tyr	Gln Met Ser Thr	His Gly Lys Val Arg	310 315	
320	Ile Phe Arg Asp Ile Ala	Ser Pro Lys Phe	Ser Met Pro Ser Pro	325 330	
335	Arg Thr Val Ile Thr	Ala Asp Phe Asp	Val 340	345	
350	Ile Phe Phe Asn Asn	Asn Ile Ala Tyr Arg	Ser Ser Ser Ala Asn	355 360	
365	Leu Phe Arg Val Ile	Arg Arg Glu His	Gly Asp Pro Leu Ile	370 375	
380	Glu Leu Asn Pro Gly	Gly Asp Ala Leu Glu	Pro Glu Gly Arg Gly	385 390	
395	Gly Gly Val Val Thr	Asp Phe Asp Gly	Thr 400	405	
410	Ile Leu Ser His Gly	Asp Gly Met Leu Asp	Leu 415	420	
425	Arg Gly Asn Gln Gly	Gly Ser Met Ala Gln	Pro Leu Ser Val Phe	430 435	
440	Ile Arg Arg Gly	Pro Asn Asn Asn	Trp Leu Arg Val Val	445 450	
455	Tyr Thr Arg Phe Gly	Ala Phe Ala Arg	Gly Ala Lys Val Val	Leu 460	465
470	Ser Gly Ala His Leu	Arg Ile Ile Asp Gly	Gly 475	480	
485	Ser Gly Tyr Leu Cys	Glu Met Glu Pro Val	Ala His Phe Gly	Leu 490	495
500	Gly Lys Asp Glu Ala	Ser Ser Val Glu Val	Thr Trp Pro Asp Gly	505	510
515	Lys Met Val Ser Arg	Asn Val Ala Ser Gly	Glu Met Asn Ser Val	520	525
530	Leu Glu Ile Leu Tyr	Pro Arg Asp Glu Asp	Thr Leu Gln Asp Pro	535	540
545	Ala Pro Leu Glu Cys	Gly Gln Gly Phe	Ser Gln Gln Glu Asn Gly	550	555
560	His Cys Met Asp Thr	Asn Glu Cys Ile Gln	Phe Pro Phe Val Cys	565	570
575	Pro Arg Asp Lys Pro	Val Cys Val Asn Thr	Tyr Gly Ser Tyr Arg	580	585
590	Cys Arg Thr Asn Lys	Lys Cys Ser Arg	Gly Tyr Glu Pro Asn Glu	595	600
605	Asp Gly Thr Ala Cys	Val Gly Trp Trp	Ser Pro Val Leu Lys Ile	610	615
620	Val Thr Pro Gln Val	Gly Lys Ser Leu	Gly Pro	625	

<210> 14  
 <211> 296  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3031062CD1

&lt;400&gt; 14

Met	Glu	Trp	Trp	Ala	Ser	Ser	Pro	Leu	Arg	Leu	Trp	Leu	Leu	
1	5							10				15		
Phe	Leu	Leu	Pro	Ser	Ala	Gln	Gly	Arg	Gln	Lys	Glu	Ser	Gly	Ser
								20		25				30
Lys	Trp	Lys	Val	Phe	Ile	Asp	Gln	Ile	Asn	Arg	Ser	Leu	Glu	Asn
								35		40				45
Tyr	Glu	Pro	Cys	Ser	Ser	Gln	Asn	Cys	Ser	Cys	Tyr	His	Gly	Val
								50		55				60
Ile	Glu	Glu	Asp	Leu	Thr	Pro	Phe	Arg	Gly	Gly	Ile	Ser	Arg	Lys
								65		70				75
Met	Met	Ala	Glu	Val	Val	Arg	Arg	Lys	Leu	Gly	Thr	His	Tyr	Gln
								80		85				90
Ile	Thr	Lys	Asn	Arg	Leu	Tyr	Arg	Glu	Asn	Asp	Cys	Met	Phe	Pro
								95		100				105
Ser	Arg	Cys	Ser	Gly	Val	Glu	His	Phe	Ile	Leu	Glu	Val	Ile	Gly
								110		115				120
Arg	Leu	Pro	Asp	Met	Glu	Met	Val	Ile	Asn	Val	Arg	Asp	Tyr	Pro
								125		130				135
Gln	Val	Pro	Lys	Trp	Met	Glu	Pro	Ala	Ile	Pro	Val	Phe	Ser	Phe
								140		145				150
Ser	Lys	Thr	Ser	Glu	Tyr	His	Asp	Ile	Met	Tyr	Pro	Ala	Trp	Thr
								155		160				165
Phe	Trp	Glu	Gly	Gly	Pro	Ala	Val	Trp	Pro	Ile	Tyr	Pro	Thr	Gly
								170		175				180
Leu	Gly	Arg	Trp	Asp	Leu	Phe	Arg	Glu	Asp	Leu	Val	Arg	Ser	Ala
								185		190				195
Ala	Gln	Trp	Pro	Trp	Lys	Lys	Lys	Asn	Ser	Thr	Ala	Tyr	Phe	Arg
								200		205				210
Gly	Ser	Arg	Thr	Ser	Pro	Glu	Arg	Asp	Pro	Leu	Ile	Leu	Leu	Ser
								215		220				225
Arg	Lys	Asn	Pro	Lys	Leu	Val	Asp	Ala	Glu	Tyr	Thr	Lys	Asn	Gln
								230		235				240
Ala	Trp	Lys	Ser	Met	Lys	Asp	Thr	Leu	Gly	Lys	Pro	Ala	Ala	Lys
								245		250				255
Asp	Val	His	Leu	Val	Asp	His	Cys	Lys	Tyr	Lys	Tyr	Leu	Phe	Asn
								260		265				270
Phe	Arg	Gly	Val	Leu	Gln	Val	Ser	Gly	Leu	Asn	Thr	Ser	Ser	Cys
								275		280				285
Val	Ala	Ile	Ile	Leu	Met	Arg	Lys	Arg	Thr	Tyr				
								290		295				

&lt;210&gt; 15

&lt;211&gt; 249

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3101617CD1

&lt;400&gt; 15

Met	Asp	Gly	Lys	Lys	Cys	Ser	Val	Trp	Met	Phe	Leu	Pro	Leu	Val
1	5							10				15		
Phe	Thr	Leu	Phe	Thr	Ser	Ala	Gly	Leu	Trp	Ile	Val	Tyr	Phe	Ile
								20		25				30
Ala	Val	Glu	Asp	Lys	Ile	Leu	Pro	Leu	Asn	Ser	Ala	Glu	Arg	
								35		40				45
Lys	Pro	Gly	Val	Lys	His	Ala	Pro	Tyr	Ile	Ser	Ile	Ala	Gly	Asp
								50		55				60
Asp	Pro	Pro	Ala	Ser	Cys	Val	Phe	Ser	Gln	Val	Met	Asn	Met	Ala
								65		70				75
Ala	Phe	Leu	Ala	Leu	Val	Val	Ala	Val	Leu	Arg	Phe	Ile	Gln	Leu
								80		85				90

Lys	Pro	Lys	Val	Leu	Asn	Pro	Trp	Leu	Asn	Ile	Ser	Gly	Leu	Val
95									100					105
Ala	Leu	Cys	Leu	Ala	Ser	Phe	Gly	Met	Thr	Leu	Leu	Gly	Asn	Phe
				110					115					120
Gln	Leu	Thr	Asn	Asp	Glu	Glu	Ile	His	Asn	Val	Gly	Thr	Ser	Leu
					125				130					135
Thr	Phe	Gly	Gly	Thr	Leu	Thr	Cys	Trp	Ile	Gln	Ala	Ala	Leu	
					140				145					150
Thr	Leu	Lys	Val	Asn	Ile	Lys	Asn	Glu	Gly	Arg	Arg	Val	Gly	Ile
					155			160						165
Pro	Arg	Val	Ile	Leu	Ser	Ala	Ser	Ile	Thr	Leu	Cys	Val	Val	Leu
					170				175					180
Tyr	Phe	Ile	Leu	Met	Ala	Gln	Ser	Ile	His	Met	Tyr	Ala	Ala	Arg
					185				190					195
Val	Gln	Trp	Gly	Leu	Val	Met	Cys	Phe	Leu	Ser	Tyr	Phe	Gly	Thr
					200				205					210
Phe	Ala	Val	Glu	Phe	Arg	His	Tyr	Arg	Tyr	Glu	Ile	Val	Cys	Ser
					215				220					225
Glu	Tyr	Gln	Glu	Asn	Phe	Leu	Ser	Phe	Ser	Glu	Ser	Leu	Ser	Glu
					230				235					240
Ala	Ser	Glu	Tyr	Gln	Thr	Asp	Gln	Val						
					245									

<210> 16  
 <211> 124  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3216178CD1

<400> 16

Met	Gly	Gly	Tyr	Leu	Lys	Thr	Arg	Pro	Trp	Thr	Leu	Gln	His	Phe
1					5					10				15
Tyr	Leu	Cys	Leu	Met	Pro	Ala	Ala	Thr	Trp	Leu	Val	Leu	Leu	
					20					25				30
Leu	Leu	Trp	Leu	Ser	Leu	Gly	Val	Lys	Thr	Gly	Ser	Cys	Ser	Gln
					35				40					45
Pro	Gln	Asn	Leu	Cys	Cys	Leu	Gly	Thr	Asp	His	His	Cys	Lys	Arg
					50				55					60
Gly	Ser	Cys	Tyr	Cys	Asp	Glu	Phe	Cys	His	Val	Ala	Pro	Asp	Cys
					65				70					75
His	Pro	Asp	His	Ser	Val	Leu	Cys	Asn	Pro	Ala	Ser	Gln	Met	Thr
					80				85					90
Lys	Met	Val	Leu	Gln	Met	Val	Leu	Arg	Met	Glu	Asn	Pro	Pro	Ser
					95				100					105
Pro	Ala	Arg	Ser	His	Leu	Asp	Trp	Met	Gln	Ser	Met	Val	Ser	Ser
					110				115					120
Leu	Gln	Val	Leu											

<210> 17  
 <211> 101  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3406803CD1

<400> 17

Met	Leu	Pro	Val	Gly	Ala	Gln	Pro	Arg	Ser	Pro	Pro	Trp	Val	Leu
1				5					10				15	
Ala	Arg	Leu	Leu	His	Pro	Arg	Gly	Pro	Ala	Ala	Thr	Ser	Leu	Val
				20					25				30	
Pro	Phe	Leu	Pro	Trp	Gly	Ser	Leu	Glu	Ser	His	Thr	Pro	Cys	Pro
					35				40				45	
Tyr	Arg	Ala	Cys	Ser	Pro	Gly	Trp	Glu	Leu	Thr	Leu	Ser	Thr	Phe
					50				55				60	
Pro	Glu	Arg	Glu	Thr	Leu	Ser	Gly	Gly	Glu	Val	Arg	Lys	Arg	Gly
					65				70				75	
Ala	Gly	Ser	Met	Val	Gly	Gly	Glu	Ser	Thr	Met	Thr	Arg	Ala	
				80				85				90		
Leu	Cys	Val	Arg	Leu	Leu	Thr	Lys	Leu	Arg	Val				
					95						100			

<210> 18  
 <211> 540  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3468066CD1

<400> 18

Met	Ala	Thr	Ser	Gly	Ala	Ala	Ser	Ala	Glu	Leu	Val	Ile	Gly	Trp
1				5					10				15	
Cys	Ile	Phe	Gly	Leu	Leu	Leu	Leu	Ala	Ile	Leu	Ala	Phe	Cys	Trp
					20				25				30	
Ile	Tyr	Val	Arg	Lys	Tyr	Gln	Ser	Arg	Arg	Glu	Ser	Glu	Val	Val
					35				40				45	
Ser	Thr	Ile	Thr	Ala	Ile	Phe	Ser	Leu	Ala	Ile	Ala	Leu	Ile	Thr
					50				55				60	
Ser	Ala	Leu	Leu	Pro	Val	Asp	Ile	Phe	Leu	Val	Ser	Tyr	Met	Lys
					65				70				75	
Asn	Gln	Asn	Gly	Thr	Phe	Lys	Asp	Trp	Ala	Asn	Ala	Asn	Val	Ser
					80				85				90	
Arg	Gln	Ile	Glu	Asp	Thr	Val	Leu	Tyr	Gly	Tyr	Tyr	Thr	Leu	Tyr
					95				100				105	
Ser	Val	Ile	Leu	Phe	Cys	Val	Phe	Phe	Trp	Ile	Pro	Phe	Val	Tyr
					110				115				120	
Phe	Tyr	Tyr	Glu	Glu	Lys	Asp	Asp	Asp	Asp	Thr	Ser	Lys	Cys	Thr
					125				130				135	
Gln	Ile	Lys	Thr	Ala	Leu	Lys	Tyr	Thr	Leu	Gly	Phe	Val	Val	Ile
					140				145				150	
Cys	Ala	Leu	Leu	Leu	Leu	Val	Gly	Ala	Phe	Val	Pro	Leu	Asn	Val
					155				160				165	
Pro	Asn	Asn	Lys	Asn	Ser	Thr	Glu	Trp	Glu	Lys	Val	Lys	Ser	Leu
					170				175				180	
Phe	Glu	Glu	Leu	Gly	Ser	Ser	His	Gly	Leu	Ala	Ala	Leu	Ser	Phe
					185				190				195	
Ser	Ile	Ser	Ser	Leu	Thr	Leu	Ile	Gly	Met	Leu	Ala	Ala	Ile	Thr
					200				205				210	
Tyr	Thr	Ala	Tyr	Gly	Met	Ser	Ala	Leu	Pro	Leu	Asn	Leu	Ile	Lys
					215				220				225	
Gly	Thr	Arg	Ser	Ala	Ala	Tyr	Glu	Arg	Leu	Glu	Asn	Thr	Glu	Asp
					230				235				240	
Ile	Glu	Glu	Val	Glu	Gln	His	Ile	Gln	Thr	Ile	Lys	Ser	Lys	Ser
					245				250				255	
Lys	Asp	Gly	Arg	Pro	Leu	Pro	Ala	Arg	Asp	Lys	Arg	Ala	Leu	Lys
					260				265				270	
Gln	Phe	Glu	Glu	Arg	Leu	Arg	Thr	Leu	Lys	Lys	Arg	Glu	Arg	His
					275				280				285	
Leu	Glu	Phe	Ile	Glu	Asn	Ser	Trp	Trp	Thr	Lys	Phe	Cys	Gly	Ala

	290	295	300											
Leu	Arg	Pro	Leu	Lys	Ile	Val	Trp	Gly	Ile	Phe	Phe	Ile	Leu	Val
				305		310								315
Ala	Leu	Leu	Phe	Val	Ile	Ser	Leu	Phe	Leu	Ser	Asn	Leu	Asp	Lys
				320		325								330
Ala	Leu	His	Ser	Ala	Gly	Ile	Asp	Ser	Gly	Phe	Ile	Ile	Phe	Gly
				335		340								345
Ala	Asn	Leu	Ser	Asn	Pro	Leu	Asn	Met	Leu	Leu	Pro	Leu	Leu	Gln
				350		355								360
Thr	Val	Phe	Pro	Leu	Asp	Tyr	Ile	Leu	Ile	Thr	Ile	Ile	Ile	Met
				365		370								375
Tyr	Phe	Ile	Phe	Thr	Ser	Met	Ala	Gly	Ile	Arg	Asn	Ile	Gly	Ile
				380		385								390
Trp	Phe	Phe	Trp	Ile	Arg	Leu	Tyr	Lys	Ile	Arg	Arg	Gly	Arg	Thr
				395		400								405
Arg	Pro	Gln	Ala	Leu	Leu	Phe	Leu	Cys	Met	Ile	Leu	Leu	Ile	
				410		415								420
Val	Leu	His	Thr	Ser	Tyr	Met	Ile	Tyr	Ser	Leu	Ala	Pro	Gln	Tyr
				425		430								435
Val	Met	Tyr	Gly	Ser	Gln	Asn	Tyr	Leu	Ile	Glu	Thr	Asn	Ile	Thr
				440		445								450
Ser	Asp	Asn	His	Lys	Gly	Asn	Ser	Thr	Leu	Ser	Val	Pro	Lys	Arg
				455		460								465
Cys	Asp	Ala	Glu	Ala	Pro	Glu	Asp	Gln	Cys	Thr	Val	Thr	Arg	Thr
				470		475								480
Tyr	Leu	Phe	Leu	His	Lys	Phe	Trp	Phe	Phe	Ser	Ala	Ala	Tyr	Tyr
				485		490								495
Phe	Gly	Asn	Trp	Ala	Phe	Leu	Gly	Val	Phe	Leu	Ile	Gly	Leu	Ile
				500		505								510
Val	Ser	Cys	Cys	Lys	Gly	Lys	Lys	Ser	Val	Ile	Glu	Gly	Val	Asp
				515		520								525
Glu	Asp	Ser	Asp	Ile	Ser	Asp	Asp	Glu	Pro	Ser	Val	Tyr	Ser	Ala
				530		535								540

<210> 19  
<211> 108  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 3592862CD1

<400> 19  
Met Thr Pro Ser Arg Leu Pro Trp Leu Leu Ser Trp Val Ser Ala  
1 5 10 15  
Thr Ala Trp Arg Ala Ala Arg Ser Pro Leu Leu Cys His Ser Leu  
20 25 30  
Arg Lys Thr Ser Ser Gln Gly Gly Lys Ser Glu Leu Val Lys  
35 40 45  
Gln Ser Leu Lys Pro Lys Leu Pro Glu Gly Arg Phe Asp Ala  
50 55 60  
Pro Glu Asp Ser His Leu Glu Lys Glu Pro Leu Glu Lys Phe Pro  
65 70 75  
Asp Asp Val Asn Pro Val Thr Lys Glu Lys Gly Gly Pro Arg Gly  
80 85 90  
Pro Glu Pro Thr Arg Tyr Gly Asp Trp Glu Arg Lys Gly Arg Cys  
95 100 105  
Ile Asp Phe

<210> 20

<211> 114  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3669422CD1

<400> 20  
 Met Ser Ser Ser Ser Arg Cys Leu Ser Pro Ser Pro Gly Met  
 1 5 10 15  
 Ser Leu Trp Ser Cys Leu Leu Phe Leu Cys Thr Pro Ser Pro Thr  
 20 25 30  
 Thr Thr Ser Pro Ser Pro Asp Pro Ser Gln Val Ser Thr Leu Pro  
 35 40 45  
 Thr Pro Ser Pro Gln Arg Glu Gly Leu Lys Gln Gly Gln Trp Arg  
 50 55 60  
 Lys Thr Gly Pro Ser Ser Thr His Pro His Thr Pro Ser Ser Arg  
 65 70 75  
 Pro Pro Ser Pro Ser Ser Leu Pro Leu Thr Trp Lys Leu Leu Gln  
 80 85 90  
 Pro Ile Pro Ser His Ser Leu Pro His Pro Pro Lys Ile His Thr  
 95 100 105  
 Gly Pro Ser Leu Ala Glu Cys Gly His  
 110

<210> 21  
 <211> 114  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3688740CD1

<400> 21  
 Met Arg Gly Glu His Asn Ser Thr Ser Tyr Asp Ser Ala Val Ile  
 1 5 10 15  
 Tyr Arg Gly Phe Trp Ala Val Leu Met Leu Leu Gly Val Val Ala  
 20 25 30  
 Val Val Ile Ala Ser Phe Leu Ile Ile Cys Ala Ala Pro Phe Ala  
 35 40 45  
 Ser His Phe Leu Tyr Lys Ala Gly Gly Ser Tyr Ile Ala Ala  
 50 55 60  
 Asp Gly Ile Ser Ser Leu Cys Tyr Ser Ser Leu Ser Lys Ser Leu  
 65 70 75  
 Leu Ser Gln Pro Leu Arg Glu Thr Ser Ser Ala Ile Asn Asp Ile  
 80 85 90  
 Ser Leu Leu Gln Ala Leu Met Pro Leu Leu Gly Trp Thr Ser His  
 95 100 105  
 Trp Thr Cys Ile Thr Val Gly Leu Tyr  
 110

<210> 22  
 <211> 287  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3742589CD1

&lt;400&gt; 22

Met	Glu	Leu	Glu	Arg	Ile	Val	Ser	Ala	Ala	Leu	Leu	Ala	Phe	Val
1					5				10				15	
Gln	Thr	His	Leu	Pro	Glu	Ala	Asp	Leu	Ser	Gly	Leu	Asp	Glu	Val
					20				25				30	
Ile	Phe	Ser	Tyr	Val	Leu	Gly	Val	Leu	Glu	Asp	Leu	Gly	Pro	Ser
					35				40				45	
Gly	Pro	Ser	Glu	Glu	Asn	Phe	Asp	Met	Glu	Ala	Phe	Thr	Glu	Met
					50				55				60	
Met	Glu	Ala	Tyr	Val	Pro	Gly	Phe	Ala	His	Ile	Pro	Arg	Gly	Thr
					65				70				75	
Ile	Gly	Asp	Met	Met	Gln	Lys	Leu	Ser	Gly	Gln	Leu	Ser	Asp	Ala
					80				85				90	
Arg	Asn	Lys	Glu	Asn	Leu	Gln	Pro	Gln	Ser	Ser	Gly	Val	Gln	Gly
					95				100				105	
Gln	Val	Pro	Ile	Ser	Pro	Glu	Pro	Leu	Gln	Arg	Pro	Glu	Met	Leu
					110				115				120	
Lys	Glu	Glu	Thr	Arg	Ser	Ser	Ala	Ala	Ala	Ala	Ala	Asp	Thr	Gln
					125				130				135	
Asp	Glu	Ala	Thr	Gly	Ala	Glu	Glu	Glu	Leu	Leu	Pro	Gly	Val	Asp
					140				145				150	
Val	Leu	Leu	Glu	Val	Phe	Pro	Thr	Cys	Ser	Val	Glu	Gln	Ala	Gln
					155				160				165	
Trp	Val	Leu	Ala	Lys	Ala	Arg	Gly	Asp	Leu	Glu	Glu	Ala	Val	Gln
					170				175				180	
Met	Leu	Val	Glu	Gly	Lys	Glu	Glu	Gly	Pro	Ala	Ala	Trp	Glu	Gly
					185				190				195	
Pro	Asn	Gln	Asp	Leu	Pro	Arg	Arg	Leu	Arg	Gly	Pro	Gln	Lys	Asp
					200				205				210	
Glu	Leu	Lys	Ser	Phe	Ile	Leu	Gln	Lys	Tyr	Met	Met	Val	Asp	Ser
					215				220				225	
Ala	Glu	Asp	Gln	Lys	Ile	His	Arg	Pro	Met	Ala	Pro	Lys	Glu	Ala
					230				235				240	
Pro	Lys	Lys	Leu	Ile	Arg	Tyr	Ile	Asp	Asn	Gln	Val	Val	Ser	Thr
					245				250				255	
Lys	Gly	Glu	Arg	Phe	Lys	Asp	Val	Arg	Asn	Pro	Glu	Ala	Glu	Glu
					260				265				270	
Met	Lys	Ala	Thr	Tyr	Ile	Asn	Leu	Lys	Pro	Ala	Arg	Lys	Tyr	Arg
					275				280				285	
Phe His														

&lt;210&gt; 23

&lt;211&gt; 854

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 078811CB1

&lt;400&gt; 23

attttgcctc	gtggacccaa	aggttagcaat	ttgaaacatg	aggagtacga	ttctactgtt	60
ttgtcttcta	ggatcaactc	ggtcattacc	agtctttct	tctttaagtc	tgataccatt	120
aacacagatg	ctcacactgg	ggccagatct	gcatctgtt	aatcctgtg	caggaatgac	180
acctggtacc	cagacccacc	cattgaccct	gggagggtt	aatgtacaac	agcaactgca	240
cccacatgtt	ttaccaattt	ttgtcacaca	acttggagcc	ccagggcact	atcctaagct	300
cagaggaatt	gccacaaatc	ttcacgagcc	tcatcatcca	ttccttgg	cccgggaggc	360
atccttgc	accagtccagg	caggggctaa	tccagatgtc	caggatggaa	gccttccagc	420
aggaggagca	ggtgtaaatc	ctgcccaccc	gggaacccca	gcaggccg	tcccaactcc	480
cagtggcaca	gatgacgact	ttgcagtgtac	cacccctgca	ggcatccaaa	ggagcacaca	540
tgccatcgag	gaagccacca	cagaatcgtc	aaatggatt	cagtaagctg	tttcaaaattt	600
tttcaactaa	gctgcctcga	atttggtgat	acatgtgaat	ctttatcatt	gattatatta	660
tggaatagat	tgagacacat	tgatagtgat	tagaagaaat	taattcttaa	tttacctgaa	720
aatattcttg	aaatttcaga	aaatatgttc	tatgttagaga	atcccaactt	ttaaaaacaa	780

taattcaatg gataaatctg tcttgaaat ataacattat gctgcctgga tgatatgcat 840  
 attaaaacga atta 854

<210> 24  
 <211> 1804  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 371156CB1

<400> 24  
 gtgataggca gcttccttc tttcaacag tgataccatc gaaaatcaa ataaatgcaa 60  
 gctgaggttt tgtgctcaact gaaaggcgt tcaacccag aaggccgaca caaaaaaaaaat 120  
 ggtatgtgaa gatgcaccgt ctttcaaat ggcctggag agtcaaatgg cctggagag 180  
 ggggcctgct cttctgtgct gtgcctttc ggcttcccgat ttgagctccc aagaccagga 240  
 cccactgggg catataaaaat ctctgtgtc tcccttcggc tttcccgatgg agctccaaag 300  
 accaggaccct actggggcat ataaaaaaagt caaaaatcaa aatcaaacaa caagttctga 360  
 gttacttagg aaacagactt cgcatatcaa tcagagaggc cacagagcaa ggtctaaact 420  
 tctggcttc agacaaattc ctgtatagaac atttaaatgt gggaaagtggc ttcccccaggt 480  
 cccatccccct gtttagggat agagttgata tcattttat aggtgccatg tatgcctctg 540  
 cctgaattttt ttaatttgac ttttaggctt ttgagatgc acgaggggaga acaaggccctt 600  
 tgctgttgg gataggaaag acttaaccta aaattaaacc acgaagaaag cattgtaaa 660  
 aatctaacaat tataaggcc ttttatggat cattttttc aaaaatggaa aactccagaa 720  
 acgcacagga acgaaatacc tccagaaac atgaagcaat catcagaac tcactggtaa 780  
 tattttaaa aagtatacag atcaaagcaa aaagaagcca tttgttacaa agagaaaatgt 840  
 gcaaataattt ttaaggccat tataagtgc aagaggagta acatgaaata aacattcttt 900  
 cacatggcta ctggaaatataat aaatttcgct ccagaaaggc cgtacgatgt tgacgatagg 960  
 tggaaaacc ttaagatgt gtactggggc ccagaatattt tatttctagg aatgtatct 1020  
 gaggaaatataa tccgagatcc ccacaaactg caatgtttag gaaattgtcct tatagcatgg 1080  
 catacacaag aaaaacagag aaaaacgtga tccctgtcag tggaaaagggttcaatgaa 1140  
 ttacgggtgt tctgcattgt gcttttatgt cattaaaaat tttgttacaa cggccaggca 1200  
 cagtggctca tgcctgtat cctaacactt tgggaggccca aggtggaaag attgcctgag 1260  
 ctcaggagtt tgagaccgc ctggcaaca cggtaaaacc cctgtctctac taaaatacaa 1320  
 aaaatttagcc gggcgtcgca gcatgcgcct gtagtcccgat ctgtcagga ggctgaggca 1380  
 ggagaattgtt ttaacccggg gaggcagagg ttgcactgtat cttggatggaa gccaccgcac 1440  
 tccagcctgtt ggcacagac aagattccgt tcccaagaaa aaaaattgtt tcaacaataa 1500  
 gggcaaaagg agagaatcat aacatctgtat taaacagaaa aagcaagattt tttaaaacta 1560  
 actatataag gatggtccca gctgtgtcaa aaggaagctt gttttaataa cgtgtgcata 1620  
 aaaattaaat agaggtgaac acaattattt taaggcagtt aaattatctc ttttttttgc 1680  
 actaagactt tctagaattt tacttattca ttctgtactt aaatttttc taatgaacac 1740  
 atataactttt gtaatcagaa aatattaaat gcatgttattt ttcaaaatca aaaaaaaaaa 1800  
 aaaa 1804

<210> 25  
 <211> 2663  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 584050CB1

<400> 25  
 ggagaaagga tggccggcct ggcggcgcgg ttggtcctgc tagctggggc agcggcgctg 60  
 gcgagcggct cccaggcgca ccgtgagccg gtgtaccgcg actgcgtact gcagtgcgaa 120  
 gagcagaact gtcgtggggg cgtctgtat cacttccgtt cccggccagcc aatctacatg 180  
 agtctagcag gctggacgtc tggggacgc ttttttttttgcgtatgtt ggttccgggtt 240  
 gggctctacc tccaggaagg tccacaaatgt ctttttttttgcgtatgtt ggttccgggtt 300  
 cgggttcctgt ttttcaaga gccggcatcg gccgtggcct ctttttttttgcgtatgtt ggttccgggtt 360  
 agcctgtgtat ttttgcgtatgtt ggttccgggtt ctttttttttgcgtatgtt ggttccgggtt 420  
 acctgtgtgg ctttcgttgc ggttccctc aatgcgtatgtt ttttttttttgcgtatgtt ggttccgggtt 480  
 accagggaca ctgacccatc agagaaaatgt gactacttgc ttttttttttgcgtatgtt ggttccgggtt 540  
 cactcaatct acctgtgtat ggttccgggtt ctttttttttgcgtatgtt ggttccgggtt 600

gccttcggg ctctccgtc gctcatgctg accgtgcacg tctccctaccc gagecctcatc 660  
 cgcttcgact atggctacaa cctggtgccc aacgtggcta ttggcctggg caacgtggg 720  
 tgggtggctg cctgggtcct gtggaaaccag cggcggtgc ctcacgtgcg caagtgcgtg 780  
 gtgggtgtct tgctgtcga ggggtgtcc ctgctcgagc tgctgtactt cccaccgctc 840  
 ttctgggtcc tggatgccc tgccatctgg cacatcgca ccatccctgt ccacgtcctc 900  
 ttttcagct ttctggaaaga tgacagcctg tacctgtca aggaatcaga ggacaagtcc 960  
 aagctggact gaagacccg gaggcgtct gccccagtt ggatccgtcc cccgcctgc 1020  
 tggcctccct tctccctca acccttgaga tgattttctc ttttcaactt cttgaacttg 1080  
 gacatgaagg atgtggggcc agaatcatgt ggccagccca cccctgttgc gcccctacca 1140  
 gccttgggt ctgttctagg gaaggcctcc ctcacacttc tggagctgaa ctgggggtggaa actgagtgt tgcttagetc taccggggagg 1200  
 acagctgcct gtttccccc caccagcctc ctcccccacat ccccaactgc ctggctgggt 1260  
 cctgaagccc tctgtetacc tggagagacca gggaccacag geettagggg tacaggggg 1320  
 ccccttctgt taccaccccc caccctcctc caggacacca cttaggtgtg ctggatgttt 1440  
 gttctttggc cagccaaggt tcaacggcat tctccccatg ggatcttgag ggaccaaggt 1500  
 gctgggattt ggaaggagtt tcaacccgtac ctggccctta gccaggttcc caggaggcct 1560  
 caccactact ccttccaggg ccagggtctc agcaagccca gggcaaggat ctgtctgc 1620  
 tgcgtgttgc agacggcttcc acctgtgtc gggaggttgg gccaggttgc gtgcataagg 1680  
 gacaggggccg tgagcatggg cctgggtgtg tgcgtgttgc ggcctagggt cgcaagtgtt 1740  
 agacgggtgt tgcggggaa gaggtgtggc ttcaaaagtgt gtgtgtgcag ggggtgggtg 1800  
 tggtagcgtg ggttagggaa acgtgtgtgc gctgtgttgc gggcatgtga gatgagtgtac 1860  
 tggccgggtt gttgtccaca gttgagaggtt tggagcagga tgagggaaatc ctgtcaccat 1920  
 caataatcac ttgtggagcg ccacgtctgc ccaagaccc acctggggcgg acagccagg 1980  
 gctctccatg gccaggctgc ctgtgtcat gttccctgtc ttgtgtccct ttgcccgcct 2040  
 cctgcaaacc tcacagggtc cccacacaaac agtgcctcc agaagcagcc cctcggagcc 2100  
 agaggaaggaa aaatggggat ggctggggct ctctccatcc tccttttctc ttgccttcg 2160  
 catggctggc ctccccctcc aaaacctcca ttccccctgt gccagccct ttgcataatc 2220  
 ctgattttgg ggaggagggaa gggcgattt gaggggagaag gggagaaaagc ttatggctgg 2280  
 gtctggtttc ttccttcccc agaggggttctt actgttccag ggtggcccca gggcaggcag 2340  
 gggccacact atgcctgcgc cctggtaaaag gtgacccttgc ccatttacca gcaagccctgg 2400  
 catgttccctt ccccacagga atagaatggaa gggagcttca gaaaccccttcc atcccaaagg 2460  
 cagtctccgt ggttggaaaga gactggattt ttgctctggc cctgaccctt tgcccttctt 2520  
 tgaggagggg gagctatgtt aggactccaa cctcaggac tcgggtggcc tgcgtactgt 2580  
 tcttttgata ctgaaaaactt ttaaggtggg agggtggcaa gggatgtgtca taataaatca 2640  
 attccaagcc tcaaaaaaaaaaaa aaa 2663

<210> 26  
 <211> 769  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 863808CB1

<400> 26  
 gcgaccccgca cgcaaggctg ctgctatggg gccggggggc cgtgtggcgc ggctgctcgc 60  
 cccactaatg tggcgcagggg cggtttcctc ggtggcgggg tcccgcggtt gagccgagcc 120  
 cgggcttcggg ctgctggccg tgcagcggct tcccgttagga gcagcggttgc gccgggcttgc 180  
 ccagacccca aactttgtcc gggccctgc cagcgagcc gggctggagg agcggggcgg 240  
 ggggacgggtt aacgaggggac gcccagaatc ggacgcggca gatcatactg gtcccaaggt 300  
 tgacatcgat atgtgtttt cacttctgag gcaagaaaaat gcaagagaca tttgtgttat 360  
 ccaggttccctt ccagaaatga gatatacaga ttactttgtt attgttagt gaaacttctac 420  
 ccgacacttta catgccatgg ctttctacgt tgcgtttttt gataacacc taaaatgtaa 480  
 acgtgaccctt catgttaaga tagaaggggaa ggacactgt gactggctgt gcgtggattt 540  
 tggcagcatg gtgattcatt tgatgtttcc agaaaccaga gaaatctatg aattagagaa 600  
 attatggacc ctacgttctt atgatgacca ttatgtttttt gtagcttccatg agacagtacc 660  
 tgaagacttc attcttggaa tagaagatga tacttcatct gtgactccag tggagttaaa 720  
 atgtgaataaa aatatttttt gcaactgcgtt agtcaaaaaaaaaaaa 769

<210> 27  
 <211> 1257  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 978433CB1

<400> 27

gaggcgcgcg	ggtgaaaaggc	gcattgatgc	agcctgcggc	ggcctcggag	cgcggcggag	60
cagacgctga	ccacgttcct	ctcctcggtc	tcctccggct	ccagctccgc	gctgcccggc	120
agccgggaggc	catgcgaccc	cagggccccg	ccgcctcccc	gcagcggctc	cgcggcctcc	180
tgctgctcct	gtctgtcgag	ctgcccgcgc	cgtcgagcgc	ctctgagatc	cccaagggga	240
agcaaaaaggc	gcagctccgg	cagagggagg	tggtgacact	gtataatggc	atgtgtttac	300
aaggccgcgc	aggagtgcct	gttcgagacg	ggagccctgg	ggccaatggc	attccgggta	360
cacccggat	cccaggctgg	gatggatca	aaggagaaaa	ggggatgt	ctgagggaaa	420
gctttggaga	gtcctggaca	cccaactaca	acgagtgttc	atggagttca	ttgaattatg	480
gcatagatct	tggaaaattt	gcccggatgt	catttacaaa	gatgcgttca	aatagtgc	540
taagagtttt	gttcagttgc	tcacttcggc	taaaatgcag	aaatgc	tgtcagcgtt	600
gttatttca	attcaatgg	gctgaatgtt	caggacctct	tcccattgaa	gctataattt	660
atttggacca	aggaagccct	gaaatgaatt	caacaattaa	tattcatcgc	acttcttcgt	720
tggaaaggact	ttgtgaaagg	attgggtctg	gatttagtgg	tgttgctatc	tgggttggca	780
cttggtcaga	ttacccaaaa	ggagatgttt	ctactggatg	gaattcagtt	tctcgcatca	840
ttatttgaaga	actacccaaa	taaatgttt	aattttcatt	tgctacccct	tttttttattt	900
tgccttggaa	tggttcactt	aaatgacatt	ttaaaaatgt	ttatgtatac	atctgaatga	960
aaagcaaaggc	taaatatgtt	tacagaccaa	agtgtgat	cacactgttt	ttaaatctag	1020
cattattcat	tttgcttcaa	tcaaaagtgg	tttcaatatt	tttttttagtt	ggttagaata	1080
ctttcttcat	agtcacattc	tctcaaccta	taatttggaa	tattttgtt	gtcttttgtt	1140
ttttctctta	gtatagcatt	tttttttttt	tataaaagct	accaatctt	gtacaattt	1200
taaatgtttaa	gaattttttt	tatatctgtt	aaataaaaat	tatttccaac	aacctta	1257

<210> 28  
 <211> 2560  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1655369CB1

<400> 28

ttccagtggaa	gagcaagtgc	tgcccgaccc	aggacccctgt	gccaggctag	cagccctcca	60
gtcccttcca	gagaggaaac	ctctgtctgg	ctgagggtgg	gactagctgg	gatgtctcac	120
tccagttgt	cagggttccacc	caggaagctc	ctccgtggag	tggccagct	gattctagcc	180
ctgtcctctc	tggcagcaca	tgcacaccc	gcctgggcct	tctgtccct	gatgtttgt	240
gagccctgtc	ctcctcaatg	tttctcaag	acagaccccc	ctgaggccag	tttgaatgt	300
aagactgctg	aagtctgt	gttcacttg	agctgcagaa	aagggtggctg	ggatggccca	360
gtgtcacc	gaggcccccag	ccctttggct	gcctttgggt	tgtactttt	tttgcgttctg	420
aggccctgtcc	agagctggc	ctgcgggtgg	tggggcggtcc	gacccctggcc	agtca	480
ccgcagccct	agca	ccatgt	ccatgt	ttgttccat	aggaaagagc	540
gcctcatttggaa	accaggagtc	ccaaaaatgtc	aggagctgg	aggctctgaa	aggagcagg	600
atccatagt	gctgttggat	gaaatagggc	cccttctggg	gagcccccag	caaaaactgtt	660
tttcatatccc	actccatggaa	ctgccccct	ccagctccag	cccccacgc	agctgggttgc	720
caggcgtcat	tggagaggcc	tgtgtccccc	aggggcacca	gggatgttgc	gacctgtatg	780
ggctggcagg	aggccatgtt	ccatgttgcac	aagtgttcc	tgcccttccat	gcctggaggcc	840
accccttcagg	tggcctgtt	gcacccctca	tccggaggtt	gcctggccca	cctgttaggc	900
gagggggtctc	ttgttggagg	cctgcacagg	aagcaatgt	agcccccgtt	ccccagatgt	960
gttccactt	agccctggcg	agatggctg	tcctgttgc	tctgtccca	gacccca	1020
tctggggagc	acagtttca	ggctgttgc	tccagggagg	gggttgcggct	ctgtcaggaa	1080
acctggactc	tcaaggccca	ccagctcc	cgtgttgc	agaaatcaca	gatacagtat	1140
atacttaatt	acactaaatt	attgttgg	tttcttataa	gcactaattt	tacctgatta	1200
taggtttaaaa	tattttttt	gtcaaaaat	tttcttggg	atgtgtttt	ccctttctgc	1260
gttcatttgtt	gtgttggatgt	gaaaacttac	catccctcc	tgcttaccc	tttggccact	1320
ggggcggcaga	gaatggcgct	atgtgcagg	ggggccctgg	caccatggc	ctttggccctg	1380
cctgttgcag	agttagccct	cctggggatgt	ctccaggac	tgagcaggcc	atctgtggcc	1440
aggctgagag	aatgacttgc	tgccttacca	gcgttgcattt	gacaaggagc	tttggagct	1500
caaggggtt	ttgttggct	ggcttagagg	gaaagggtac	catccgtctg	tcctctgtc	1560
tttcttattt	cgccctccat	tgagtgttgc	tgccttgg	caactagc	ccccaccac	1620
cccaccatgc	cacccatgtt	tcttggggcc	tgtgttgc	ctccagcccc	tggggaggag	1680
aggacccagc	ccggagatgt	ggggcaaggg	cccaacatg	cccaagg	acagatgtc	1740

gcagggcagc tgctgccat gctcacgctc ctgccccct ccttcccgt gccacacccc 1800  
 accctggcc cccgcagaca cgcacatctca actcagttg gcccagcctt ctggatggct 1860  
 tgggttaggc catggggcca cctggggcca ggccagcccc tggggcagct ctggaaagagc 1920  
 agtgtggagg agcacttgct tgcaagctgg cttcagcctc tggactgtct ggagtgggcc 1980  
 ctggagctt ctgcactgtc ggctttggg acgtctcacc cacttgggtt acagttaggcc 2040  
 ttccccaccc agagagaagt gttccaccc cagagacatt gtctgtcagc ccctgaagt 2100  
 ctgcctccc ccagtggcc tcaccagccc ttccatctg tgggtccaa gtcaggctc 2160  
 ccctggggcc accagccata gggagcagcc atcagccccc gagtccagaac tgcttctgtc 2220  
 tgcctatacc tccaggtctc cccggaggg gggacggata ttatccctt aaagtttgc 2280  
 ctaatttgtt aggatttctca ggattgttgg gggctactga aaagaggaat gtgttgaatg 2340  
 tcgcgttgc tgcctactcg tccatggat ttatgttt tgctactgtc atgtgtttct 2400  
 gtgggcagag ctgggtctgg aggggtggc agtgcaccc aggtccagag catccatcca 2460  
 ccccaactggc ctccttcca gataccctct ctctaatgg gttcttgcatt gtaaaaataact 2520  
 ccacaataaa taaaataattt aacaaattaa aaaaaaaaaa 2560

<210> 29  
 <211> 614  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1703244CB1

<400> 29  
 gtgcaagagg aacaaagaaa gggactcctg cagcgtccgg ctgacctggc cttgtcata 60  
 tatctcatcc ttgctggctt cttcaactctg ttccggggcc tgggtgtct tgattgcccc 120  
 acagatgcct gctttgtcta tatctaccag tatgagccat acctgcggga ccctgtggcc 180  
 taccctaagg tgcagatgtct gatgtacatc ttttatgtcc tgcctttctg cggcctggct 240  
 gcctatgtct tcacccccc ttgttgcctc tggctccag actggccctt ggtgtttgt 300  
 ggaggcatcg gcccggcaaa gtttcgcac atggggctt ccatgcaccc ggcacacccc 360  
 ttcacctacc gtgtgcctga ggacacctgg ggctgtttt tctgtgtcaa tctgtgttat 420  
 ggcgtggcc cccacccgtt ggcttaccgt tgccttcaactt ggccgcatt cttccaccag 480  
 ccaccacccct ccgacccccc agccctccac aagaagcagc attgagagag ctgtggactc 540  
 aggacccagg actctgttta ctgtccctgt cagccctacc tggggaaagcg ggggttgggt 600  
 gtttagaga cagg 614

<210> 30  
 <211> 1936  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1730819CB1

<400> 30  
 gactacgggc tcacagccgt cccttcgtg gtggaaagaa gcccggatgg cggcagccag 60  
 cgctggggca acccggtgtc tctgtcttt gctgtatggc gtagcagcgc ccagtcgagc 120  
 cccggggcagc ggctggccgg cccggactgg tgcgcgggg gctggggccg aaggtcgaga 180  
 gggcgaggcc tggggcactgg tggggctgtct gctggggcac tcatttgaga tcgtatgcac 240  
 tgccaacttc cggaaaggccc gctcaactgtct ctggaaaccag caggatggta cttgtccct 300  
 gtcacagcgg cagctcagcg aggaggagcg gggccgactc cgggatggc cagccctgaa 360  
 tggcctgtac cgggtccggta tcccaaggcg accccggggcc ctggatggcc tggaaagctgg 420  
 tggctatgtc tcttcctttc tccctgcgtc ctccctgtt gaggcgcacc tgctggacca 480  
 gctgaccctgtc cacgtggatg tggccggcaa cgtggggccg gtgtcggtgg tgacgcaccc 540  
 cggggggctgc cggggccatgg aggtggagga cgtggaccc tggcttca acacccctgt 600  
 gcagctgcag cggcccccacca cagccccagg ccctgagacg gcccggcccttca ttgagcgct 660  
 ggagatggaa caggccccaga aggccaagaa ccccccaggag cagaagtccct tcttcgcctaa 720  
 atactggatg tacatcatcc ccgtcgtctt gtcttcgtatg atgtcaggag cggccagacac 780  
 cggggggccag ggtgggggttgg ggggttgg tgggtgggg ggttagtggcc ggtgaggggcc 840  
 caggctggtc agcgtccctgt ctgcacacc caggggccctc ctttctgtct ggagtccct 900  
 gtgtccctcaag ccattccaaag aagggtttgc tggccctcc tttcccccgg tcccacggg 960  
 ccacctgggc cagcccccttgc tccctgtctt tctgtgtggca gaggagcagc tggactgggg 1020  
 ctttggcact agcagccggc gtcttcgtc cccggccctcc ccatggccccc atgcagcccc 1080

aggggcttcc cccctgcccc tggagttagag cccgagatcc tggccactat gccagttctg 1140  
 acctcgatc cccctacccc gagcccatgc agtctggaa catggccct tctctccagc 1200  
 ctctgtgcct ttgttccagg tggctcacc ctccctgtccc tggctggct aggtggtcct 1260  
 gtccaggctc ctgcagggcc cccctcactt tgacactgga cttaggatca gcctccctc 1320  
 tgggtccccc tgagggtacc ctgggtcccc tcatacaggg cagaggcatg aaagagtccg 1380  
 ggctggatgg ccgggggctt ctggggccga cgccctagtc agccctggg gtcgtggtt 1440  
 gacatttgc tgcctgggtc aaacaaggaa tccttgccct taaggtgaca ggcctccac 1500  
 aggcttccag acttgaagga aaaggttaa gaaagaaaac aaaaccaaca gttagtgag 1560  
 tcaaagccca gacactgtaa atagaacccc ctccaccacc ccccccggcc cagcatccct 1620  
 cctggactgc ggtgctacga gggcctgccc gcctttgtc tggccaccc tccctgtaaag 1680  
 tctatttaaa aacatcgacg atacattgaa atgtgtaaac gttttgaaaa gctacagctt 1740  
 ccagcagcca aaagcaactg ttgtttggc aagacggtcc tgatgtacaa gcttgattga 1800  
 aattcaactgc tcacttgata cgttattcag aaacccaagg aatggctgc cccatcctca 1860  
 tggctgtg tggagctcag ctgtgttgc tggcagttt ttaaactgtc cccagatcg 1920  
 acacgcaaaa aaaaaaa 1936

<210> 31  
 <211> 1958  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1757161CB1

<400> 31  
 gccgcgcctt cagctacggc ccqagcggc cccgcgcgc cggggccggc cacagcctgc 60  
 agcggagccc acgagaggca ggcgcattggc ggagcagacc tactctgtggg cctattccct 120  
 ggtggattcc agtcaagtgt ctacatttct gatttccatt ctctttatag tctatggtag 180  
 tttcagggtcc ctaataggaa actttgaaaa tcaagataag gagaaggaca gtaatagttc 240  
 ttctgggttc ttcaatggca acagcacca taatagttc caaacaattt actctaccca 300  
 ggtctgttc cttcaatttgc gaggcatgt ctcttttta gtaatgttct ttttttttga 360  
 ctcagttcaa gtagtttta caatatgtac agcagttttt gcaacgatag cttttgttt 420  
 tcttccttc cccatgtgcc agtattttaa aagaccctgc tcacccctaga acaagatttc 480  
 ctttgggttc tggacgtt tcactgtgc tgagttgtc tcattctctc tgcgtgtcat 540  
 gtcgttcctt atctgggtt tcactggcca ttggcttc atggatgcac tggccatggg 600  
 cctctgtgtc gccatgtatcg cttttgtccg cctggccggc ctcaaggatct cttttgtgt 660  
 ttcttcagggg cttctcatct atgatgtttt tgggttattt ttctcagatct acatttcaa 720  
 tagcaacgtc atggtaagg tgccactca gccggctgtc aatcccctt acgttctatc 780  
 ccggaaagctc cacctggggc ccaatgttgg gctgtatgtt cttccctgt ctctgtctgg 840  
 aaaactggtc ttcccaagct ccactggcag ccacttcctt atgttggca tcggagacat 900  
 cgttatgcctt ggtcttcctac tatgtttttt ctttcgtat gacaactaca aaaagcaagc 960  
 cagttggggac ttctgtgggg cccctggacc tgcccaacatc tccggggcga tgcagaaggt 1020  
 ctctactttt cactgcaccc tcatcgatata cttttgttgc ctgtctactg ctactgtggc 1080  
 gtctcgattt caccggggc cccagggccg ctttctctat ttgttgcattt ttactttatt 1140  
 gcaacttcctt acgtatggctt attttaaaggg cgacctccgg cggatgtgtt ttttttttttgc 1200  
 ccactccaag tccagcagct cccgatttctt ggaagttatgtt tggatcacgt gaaaagtgtac 1260  
 cagatggccg tcatagtctt ttctctcaat ctcattttttt gttttctttt agagctggcc 1320  
 tggtaacttcag aaatgtaccc ttgtttttaagg aactggctgt tgactggatt tggcattttaa 1380  
 agggagctcg ttgcaggag agaggatgtt gggccctttt tggttccctt ttttttttttgc 1440  
 gatgttagagg tggggccctt tccaagaggg acaggccctt ccccaagcgcg ctttccctccc 1500  
 acgtttttat ggtatctgcac cagactgtt cttttttttt gatggaga ttttttttttgc 1560  
 taaaaactgaa aacacgctgg gactttttt agaactttt aactaaaaaa gatggaaaaaa 1620  
 aatggatcaaa ccgaagttt tcataatggcc ctttcggatcc ttttttttttgc gttttctatg 1680  
 ggggacttcag tttcagagaa ctgagacaga agcttttttgc tcgttattttt ttttttttttgc 1740  
 ttttttttttgc ttattttttt ttttttttttgc ttttttttttgc ttttttttttgc 1800  
 tgagttacttcag ttacaacat ccacataaga attttttttttgc ttttttttttgc 1860  
 atgttgcgtt ggcagtttgc aaggggccgg ttttttttttgc ttttttttttgc 1920  
 aacacacacaca tagatataaaaaaa aaaaatgtt 1958

<210> 32  
 <211> 1424  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1976095CB1

<400> 32

gtcaagttagg	agacaaggag	caaagtccct	tcacagcggg	aggggacgcc	agcgcctgca	60
gaggctgaggc	agggaaaaag	ccagtcccc	agcggaaagca	cagctcagag	ctggctcgcc	120
atggacatcc	tggteccact	cctgcagctg	ctgggtctgc	ttcttaccct	gccccctgcac	180
ctcatggctc	tgctgggctg	ctggcagccc	ctgtgcaaaa	gctacttccc	ctacctgatg	240
gccgtgctga	ctcccaagag	caaccggcaag	atggagagca	agaaacggga	gctcttcagc	300
cagataaaagg	ggcttacagg	agcctccggg	aaagtggccc	tactggagct	gggctgcgga	360
accggagcca	actttcagg	cttacccaccg	ggctgcaggg	tcacctgcct	agaccctaaat	420
ccccactttt	agaagttct	gacaaagagc	atggctgaga	acaggcacct	ccaatatgag	480
cggtttgg	tggctcttgg	agaggacatg	agacagctgg	ctgatggctc	catggatgtg	540
gtggctgca	ctctgggtct	gtgtctgtg	cagagccaa	ggaagggtct	gcaggaggtc	600
cggagagtac	tgagaccggg	aggtgtgtctc	tttttctggg	agcatatgtgc	agaaccatat	660
ggaagctggg	ccttcatgt	gcagcaagg	ttcgagccca	cctggaaaca	cattggggat	720
gctgtctgg	tcaccagaga	gacctggaaag	gatcttggaa	acgcccagg	ctccgaaatc	780
caaattggaaac	gacagcccc	tccttggaaag	tggctacctg	ttggccccca	catcatggga	840
aaggctgtca	aataatctt	cccaagctcc	aaggcacatc	tttgccttctt	ccccagccctc	900
caattagaac	aagccaccca	ccagccatc	tatcttccac	tgagaggggac	ctagcagaat	960
gagagaagac	attcatgtac	cacccctctag	ttccctcttc	cccaacctct	gccagggcaa	1020
tctctaactt	caatcccgc	ttcgacagt	aaaaagctct	acttctacgc	tgacccagg	1080
aggaaacact	aggacccctgt	tgatcttca	actgcaagg	tctggactag	tctcccaacg	1140
tttgctctcc	aatgttgtcc	cttccttgc	ttcccatgtt	aaagctcctc	tcgctttct	1200
cctgaggct	caccatgcg	tctcttagaa	ctggtcacaa	aagtcatgtt	gcctgcattc	1260
ctgccaagcc	ccctgaccc	tctctcccca	ctaccacctt	tttcttgagc	tgggggcacc	1320
agggagaatc	agagatgtc	gggatgccag	agcaagactc	aaagaggcag	aggtttgtt	1380
ctcaaataatt	tttaataaaa	tagacgaaac	cacgaaaaaa	aaaa		1424

<210> 33  
 <211> 2238  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2169991CB1

<400> 33

cctgtctgt	cacacccac	ggcaaggggcc	agccctgtttc	ctcccggtca	cctccaaatc	60
ttgtctgttt	taattcaact	cagaggtatg	caacttgggt	aggaggccag	gggaagtggg	120
gatggcagga	catggatggc	cccttggggca	ttggctctgg	gtgtcatggg	ctgtgagagt	180
caagaaggggc	agtggctgc	ctgacttggg	ttcgaagg	tcaacttgc	cactgcgg	240
agaactgaaa	ctagccaggc	caatttccca	ttgttctgc	tcttccagg	aggagaatat	300
tagatttagt	ctggagtcg	aaaggtagcag	gggctggggg	ttgcaggggg	atgttgagaa	360
gaagtggtcc	tttgggtcagg	gtgggaagcc	aacaggattt	cctggtgcat	ttgaggtgaa	420
agggaaagagg	cgctgtgtt	tgtgtttgg	gcctgagcag	ccagaagccg	ttgcccac	480
ctatgacaag	ggagacagcg	cctggggcag	agcccaagt	tgggtgcatt	cagggttagat	540
caggaggat	gagtgggcag	tggcggagcc	aggctggggg	tttggggg	ggcctgggtt	600
gcagggttaa	gcgtgggacc	cacgtcagat	tgggtgggg	tcatgtc	tgggggtggct	660
ctcaggctt	ccctgtggc	ttcccttcc	cagggtctt	ctcccccag	atcccttagc	720
tgggtgaggg	gcaggacaga	gccttttcc	agggaaagccc	ggcacccccct	gtgtccagg	780
gaagggggagt	ccttctagcc	cctgacagct	tctctcccc	ccccctggcc	tccccagggc	840
tagccagggt	tgagttctca	cccacctgt	ccgcctgtcc	tttttacctg	gaagcacagc	900
cttggggact	gagcaggccc	tcactgtc	ttaagaagg	gaatcagcc	ctttgtgtc	960
accacccctg	gggaagggtt	gagaggagag	aaggaagtgg	ctgtttggct	gtgacaaca	1020
tgaagactc	ctgcgtatg	aacagaggc	cagggtccgg	ccctgcagcc	cccagaacact	1080
ggactggagg	ggggccatggg	gcacccggacc	ctggctctgc	cctgggtgt	gtgacccctg	1140
tgtgtcactg	cggggacccc	ggaggtgtgg	gttcaagt	ggatggaggc	caccgagctc	1200
tctgtcttca	ccatccgtt	tgggttctg	gggtctggct	ccatctccct	gtgactgt	1260
agctgggggg	gcccccaacgg	tgtctgggggg	accacgtgg	ctgtgttgc	cccagaacgt	1320
ggcatccggc	aatggggccc	tgctcgccag	gcccgtgg	aaacccagag	cagcatctc	1380
ctcatcttgg	aaggctctgg	ggccagccg	ccctgcgcca	acaccaccc	ctgctgcaag	1440
tttgcgtct	tccctgggg	ctcctgggg	gcctgtggg	gcctccggcc	cagctcagac	1500
ccagggtct	ctgccccggcc	gactcctg	cccatctc	ggcagaccc	ggccgggatc	1560

ttgggggtct	caggagtct	cctcttggc	tgtgtctacc	tccttcatct	gctgcgccga	1620
cataaggcacc	gccctcccc	taggtccag	ccgtccccca	ccagccccca	ggcaccggaga	1680
gcacgagcat	gggcaccaag	ccagggctcc	caggctgtc	ttcacgtccc	ttatgcccact	1740
atcaacacca	gctgcccccc	agctacttt	gacacagctc	accccatgg	ggggccgtcc	1800
ttgtggcg	caactccac	ccacgctgca	caccggcccc	agggccctgc	cgccctggggc	1860
tccacacca	tccctgcacg	tggcagctt	gtctctttt	agaatggact	ctacgctcag	1920
gcaggggaga	ggcctctca	cactggtccc	ggcctca	ttttccctga	ccctcggggg	1980
cccaggggca	tggaaggacc	cttaggat	cgatgagaga	gaccatgagg	ccactgggt	2040
ttccccctcc	cagggctct	gggtgtcacc	cccttactt	aatttttggg	cctccaaataa	2100
gtgtccatca	ggtgtctggc	caggcccacc	tgctgcggat	gtggtctctg	tgtgtcggt	2160
tgtgggcaca	ggtgtgagtg	tgtgagtgac	agttacccca	tttca	ttctgctgc	2220
aactaagtca	gcaacgccc					2238

<210> 34  
 <211> 536  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2616827CB1

<400> 34  
 gcatgaactt gggggtcagc atgctgagga tcctcttccct cctggatgtaa ggaggagctc 60  
 aagtgtggc aacaggcaag accctgggg ctgaatata tttcaagtac gccctcatcg 120  
 ggactgtgt ggggtgtcgcc atatctgtg gcttccttggc cctgaagatc tgcatgatca 180  
 ggaggcactt atttgacgac gactttccg acctgaaaag cacgcctggg ggcctcagt 240  
 acaccatccc gctaaaagaag agagcccaa ggcgaaacca caatttctcc aaaagagatg 300  
 cacaggtatg tgagctgtat gtgagcgtg acgtgaagag gggttcttagc cccgtggaaa 360  
 acagccatgt gttAACatct caggatgttc tgcattcaaa caccctaaaggc tggtaatgaa 420  
 ctttcacatg gactgaatat tggaggcaaa taatagaagg aatagaatat acagtgcctc 480  
 tgcctgaag gaaaatatca tgccttctt ggaagaaacg gactgcacag aggaag 536

<210> 35  
 <211> 2177  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2991370CB1

<400> 35  
 cgggaggctc gaggcagcc cgggaccggg gctgggagca agcaggcggc ggcgcggcg 60  
 gcagaggcgg cagcgagcgc ccgttccca cgcctttagg cggcggggcc gagagcggga 120  
 gatggctcc gagcgctgac cccggcatgt ccaggatgtt accgttctg ctgtgtct 180  
 gtttctgcc catcaactgat ggtcccagc ggcgtgaacc catgttcaact gcaactacca 240  
 actcaactt ccgttccgtac tatgacgatc atcccacca gtcactat ggtgtggcag 300  
 ttactgtatg gggactttt agatcgctgt ggcggggtagc aatggaccga 360  
 acctgggttctt gaagtatgac cgggcccaga agcggctgtt gaacatcgcc gtcgtatgac 420  
 gcagtcacc ctactacgac ctgcgggacc ggcaggggaa cgccatggg gtcacagct 480  
 ggcacatcga cggggacggc cgggaggaga tctactttt caacaccaaat aatgccttct 540  
 cgggggtggc caacgtacacc gacaaggatgt tcaagttccg caataaccgg tggaaagaca 600  
 tcctgagcga tgaggtaac gtggccctgt gtgtggccag cctcttttgc ggacgctct 660  
 tggctgtgt ggacagaaag ggctctggac gctactctat ctacatttgc aattacgcct 720  
 acggtaatgt gggccctgtat gccccttattt aaatggaccc tgaggccatg gacctctccc 780  
 ggggcatttcc ggcgtcaga gatgtggctg ctgaggctgg ggtcagcaaa tatacagggg 840  
 gcccggcggt cagcggtggc cccatctca gcaaggatgtt ctcggatatac ttctgcgaca 900  
 atgagaatgg gcctaatttc cttttccaca accggggca tggcacctt gtggacgctg 960  
 cggccatgtc tgggtggac gaccccccacc agcatggcg aggtgtcgcc ctggctgact 1020  
 tcaaccgtga tggcaagtg gacatgtct atggcaactg gaatggcccc caccgcctct 1080  
 atctgcaat gacccat gggaaagggtcc gttccggga catgcctca cccaaat 1140  
 ccatgcctc ccctgtccgc acggatca cccggactt tgacaatgac caggagctgg 1200  
 agatcttctt caacaacatt gcccatttgc gctcctcagc caaccgcctc ttccgcgtca 1260  
 tccgttagaga gcacggagac cccctcatcg aggagctca tcccgccgac gccttggagc 1320

ctgaggggccg	ggcacacggg	ggtgtggta	ccgacttcga	cgagacggg	atgctggacc	1380
tcatcttgc	ccatggagag	tccatggctc	agccgctgtc	cgtcttccgg	ggcaatcagg	1440
gcttcaacaa	caactggctg	cgagtggtc	cacgcacccg	gtttggggcc	tttgccaggg	1500
gagcttaagg	cgtgtctac	accaagaaga	gtggggccca	cctgaggatc	atcgacgggg	1560
gctcaggcta	cctgtgttag	atggagcccg	tggcacactt	ttggcttgggg	aaggatagaag	1620
ccagcagtgt	ggaggtgcg	tggccagatg	gcaagatgtt	gagccggaaac	gtggccagcg	1680
gggagatgaa	ctcagtgtct	gagatcctct	accccccggg	tgaggacaca	cttcaggacc	1740
cagccccact	ggagtggtgc	caaggattct	cccagcaggaa	aatggccat	tgcatggaca	1800
ccaatgaatg	catccaggta	ccattctgtgt	gccctcgaga	caagcccgta	tgtgtcaaca	1860
cctatggaaag	ctacagggtgc	cggaccaaca	agaagtgcag	tcggggctac	gagcccaacg	1920
aggatggcac	agcctgcgtg	ggctgtgtgga	gcccctgttt	gaagatagtgt	acaccacaag	1980
tttggaaagag	ccttggtccc	tgaatcaactg	aatcaactgcc	ttgaatcacc	gcctggaaaata	2040
ccttgttgc	aggaacactt	accttggaaat	tcaactgagca	ggatacaaac	ttcttattgtt	2100
ttaagcttatt	aatacattaa	gattttgggg	tgctacaccta	cataataaaat	tcccatttcc	2160
tcttgaaaaa	aaaaaaa					2177

<210> 36  
<211> 2043  
<212> DNA  
<213> *Homo sapiens*

<220>  
<221> misc\_feature  
<223> Incyte ID No: 3031062CB1

<400> 36	cgccacgacg	cagcggggaa	tctgcagtag	gtctgccggc	gatggagtgg	tgggctagct	60
cgccgcgttcg	gctctggctg	aaaatggaaa	ctgtttttcc	tcctgcccctc	agcgcagggc	cgccagaagg	120
agtcaaggtc	aactatggaa	aatattttatg	accaaattaa	caggcttttgc	gagaattacg	180	
aaccatgttc	aagtcaaaaac	tgcaagctgt	accatgggt	catagaagag	gatctaactc	240	
ctttccgagg	aggatctcc	agaagatgtg	tggcagagggt	agttagacgg	aacttaggaa	300	
cccactatca	gatcaactaa	aacagactgt	accggggaaa	tgactgcatt	ttccccctaa	360	
ggtgttagtgg	tgttgagcac	tttattttgg	aagtgtatcg	gcgtctccct	gacatggaga	420	
tggtgatcaa	tgtacgagat	tatcctcagg	ttcctaaatg	gatggagcc	gccatcccag	480	
tcttcctcctt	cagtaagaca	tcagagtacc	atgatatcat	gtatccctgc	tggacatttt	540	
gggaagggggg	acctgtgtt	tggccaattt	atccctacagg	tcttggacgg	tgggacctt	600	
tcagagaaga	tctgtttaagg	tcagcagcac	agtggccatg	gaaaaagaaa	aactctacag	660	
catatcccg	aggatcaagg	aacaatcccg	aacgagatcc	tcttattttt	tctgtctcgga	720	
aaaacccaaa	acttggtgat	gcagaataaca	ccaaaaaaacca	gccttgaaa	tctatggaaag	780	
atacctttagg	aaagccagct	gctaaggatg	tccatcttgc	ggatcactgc	aaatacaagt	840	
atctgtttaa	ttttcgaggc	gtactgcaag	tttccggttt	aaacacctct	tcttgtgtgg	900	
ccattattct	aatgagaaaag	agaacataact	agatggaaa	tttcttaagg	gcagggaaagtc	960	
atgttattac	agatgtttat	tgatgtacctg	ttatataatca	ggactaaagct	gctgggatgt	1020	
actagaaaat	acgatatttg	tcaactgtgt	tccatggaaa	tttttagacta	gcataatacc	1080	
tggacatcat	tagatggaaa	aaaatacatg	tggtttttagt	ggtttggaaaatt	taagcattt	1140	
caagttatta	tactttagac	tcacaaattt	acccttcacc	gaattaaagcc	aaaagactt	1200	
tatccagaaa	atagttgagt	agagatgaat	acataggatc	ctgcatttgcatt	taactttctc	1260	
tttgcatttg	ggataataca	ctcaacacat	ttctcttacca	cattttaaaat	ttttattttat	1320	
atccgcctt	cttccaaaaa	ggatttaaag	cagtttacaa	aatgtatag	cagggcggga	1380	
taaaaataag	aaacagacta	tacaacatgc	atactgtgaa	gtctcttca	cttccaggtt	1440	
ccatcatatg	gaacccttat	ccttccacca	caaaaatgtat	attctctgtt	atatctgtac	1500	
ctgtgatctc	tcttttttca	aaaacgttaga	cctttttctt	tgaaaaaaa	acaatcccc	1560	
cacatataatt	ctgatggca	tcctcttca	ttttgtctgg	gactttggct	ctgcccactg	1620	
tcctcatcat	ccttgtatcc	aggaccacag	cctttgtcat	gcagacatat	tcagtttctt	1680	
cccatctccc	caacctccag	aaaaagaaaa	cccttcattt	tcacattttcc	cttcagccga	1740	
gattgaagat	gaagatgagg	atgagttcaa	ggatgaagac	caggatgagg	acaaggatga	1800	
ggatggagtc	tagatgcctc	cagagcttgg	agaggaggcc	tcggtcagcc	actccgttga	1860	
cgtggggccac	ggtgaccacac	catgaagtcc	ccactatggca	ctcgatttcc	tgtctgttca	1920	
gagttgtcgc	acatcacacc	accccctgcc	aagagcagg	gtcaccacacg	gctgaatgcc	1980	
cacggaggac	tctgttgcaga	ctctcaaggg	agccagtgaa	agaaatagaa	ataaagctgt	2040	
tgc						2043	

<210> 37  
<211> 1743  
<212> DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3101617CB1

&lt;400&gt; 37

cagcagggtca cagccctcg aggcgacagc ggccccggc caccagagca gtggtacagg 60  
 catggatggg aagaaatgca gctgtatggat gttcctaccc tttgtatcca ctttgcattac 120  
 ttcagctggg ttgtggatag tatacttcat agctgtggaa gatgacaaaaa ttttaccatt 180  
 aaattcagct gaaaggaaac ctgggtgaa gcatgcacca tatataagca ttgcagggtga 240  
 tgatccctct gcaagctgt tggttagtca agttatgaac atggcagcct tccttagccct 300  
 tgggttagct gttctgcgt tcatacaact gaaaccgaag gttttaaacc cgtggctgaa 360  
 tattagtggg ttggtggtc tgggtctggc ttccctcgga atgacccctac ttggtaattt 420  
 tcagctcaca aatgatgaag aaatccataa cgtcggaact tccttgaccc ttggattttg 480  
 cacattgacc tgctggatcc aggctgcgcgt gacactcaag gtcaacatca agaatgaagg 540  
 acggagagt ggaattccac gggttattct gtcggcatct atcactctct gtgtggtcct 600  
 ctacttcatc ctcatggccc aaagcatcca catgtatgca gccagggtcc agtggggcct 660  
 ggtcatgtc ttctgtctt atttggcac ttggccctg gatgtccggc attaccgcta 720  
 tgagattgtt tgctctgagt accaggagaa aaccatcaat ttcccttgc ttggtaggtg 780  
 agcttctgaa ttcagactg accaggtgt aacactcaca ggacttgaca tagaacctca 840  
 ggtgtgacag tgggggaggg gccagtagga cacactcaca 900  
 tttcacacac acacacacac acacattcat ggcacacattt gccaaatgag ctttcaggg 960  
 cgagtttattt cttaatgaa aaagcacaag cccttatgtg tccaaatatac cgcgttaca 1020  
 ctgaaaatattt atgcacgaca gagcaagaag cttgtcatgt atcacttctt atccgtcccc 1080  
 ttcccagcac tcccttcet tccatttc tccacatgtc tcaagcaccc taccgagtag 1140  
 ggcaggccaa atgttcctt gggataatgc caactcccgaa ctttgccttc aggtccaaag 1200  
 ggcttggaaac cagctcgta ggaagttctg aatctggcac taatattttt gagggtgataa 1260  
 tagtgtatca tagaatagga cggaaattgtt attgagatgt gaccctgtgt cgcctgtgaa 1320  
 aaggcatagt gagaagaact ttcccacgaa agccccccctc atcgttgc agtggtcggc 1380  
 tgggtggatcc cccaggagaga catatgccc acactgtgag agcggaaagccc gccgctgtga 1440  
 tctggactt atgcactgtg actgagaatg atttccaaat gtgaatatgtt gtagggacgt 1500  
 ggtctatca ggcctggaaaca agatggggc agtgaagta tggtttagtgg tttgcttca 1560  
 tagtatgcca tggataatgtt tttatatttca atagtttttca ttaagtaactt accatgatgc 1620  
 tctctaagcc tcatggacaa agatgttagac caaatgcac agctgagctt gctttgggtt 1680  
 caaccatgat caaaagaaaaa ctgagggtcac ctgcaggctt acgtgggaag ctaagacaat 1740  
 atc 1743

&lt;210&gt; 38

&lt;211&gt; 1306

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3216178CB1

&lt;400&gt; 38

ctgcaaaatgtt cctgtgagcg ctgtcatatgtt gtcactctgg tttttcagat tcttccctgt 60  
 gaggctggag ttccaggatgt gtcaaaatgtt cctctgttgc ggttagctat ttcaaggcagc 120  
 tgggatcatctt gtgtcactcc tgggtctgc cagtactgc ccagggtct gctgggtctt 180  
 ccccccaggatggggaaac cagggtggct ggctggatgt ggtggatatt taaagaccag 240  
 gccttggaccc ctgcagcaact tctatctctg ctgtatgttgc gtcacatgtt ggctggctt 300  
 cttccctctt ctgtggctgtt gcttgggtt gaaagacaggc agtgcctccc aaccccaagaa 360  
 ctttgcgtt ctggggacgg atcaccactg caagagggga agttgtactt gtgtatgtt 420  
 ctgcatgtt gcaccagactt gcaacccaga ccacagtgtc ctctgtcaacc ctgcttctca 480  
 gatgaccaag atgggtgtgc agatgggtgtt gaggatggag aaccccaacaa gccccgttag 540  
 gagccacacta gactggatgc agagcatgtt gagtccctgtt cagggttctt gagaaggggt 600  
 ggtggcaggc ctgtcttgc ctttgcgtt cttccaggcccc caaaggcagg gaacccaaag 660  
 aaaaatgggggg cctgtatgtt ggcagagctc cactgcataatgtt gggtagggac 720  
 caggattgtt gtcgtggac actgaaatgtt gaaatctata tactttgtt acaaaaacatt 780  
 ctccctctt tggatgttgc ttcaccatctt aaaaatgttgc aacattttt acgcctatgtt 840  
 ttgcacaaaaa acaggcagtg gccagatgtt gcccataatgtt gatgttttgc tgacttctc 900  
 cttaaatcat cttccatttc tttccctctt gttccttgc tttgtgtt actgacaaag ccactttccc 960  
 taaaatgggggg tttccctctt tttgtgttgc tgaagccat atgcaaaaacc gaaagtgtac 1020  
 cttcaaggcactt gcaaggcttta tttgtgttgc atgaaatgtt gaaagtgtac gttggctcac 1080

aaatcaactt ttctgctcgt gagacccagg aagtacacaga tacagggcat cttagtgaa 1140  
 gggctgagc attaaaagca agggaggag tggccagggt caatggctca cttccataaa 1200  
 cccagaactt tgggaggcca aaatgagagg attgctgaga ccaggagttc gagaccatcc 1260  
 tggtaacat agtgatacac cccatctct aaaaaataa aaatga 1306

<210> 39  
 <211> 851  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3406803CB1

<400> 39  
 gggctggcca cactgcaggc gctgcaggaa gcaaaggatg aaactgtatct tttcaactaac 60  
 cagcataggc cacaggccaa agaaaacttt gttaactctc ttgtgagcca gtgtaaagtt 120  
 gacttggctt ctttggaa ggttcttc tgcagaaata aaaacacttg tctgaaagag 180  
 gctgtacta tctgtggctt gttggcaag gacacttcag atactggctt tgagctcact 240  
 gggctctggc ccctcatggc tctgtatca tctctgtccc acattgcagg ccattgtccc 300  
 ttttgggtct cagcccaggaa gccccttgc ggtcctggcc agactcctcc accctcggt 360  
 gcctgcccga accagtctag tgcctttctt cccatggggg tcccttggat ctccacacacc 420  
 ctggcccttac agagcctgca gtccagggtg ggagtttaacc ctttccactt tcccagagag 480  
 ggagaccttg agtggggggg aggtcaggaa acgcggggctt ggcagcatgg tgggaggagg 540  
 taaaagtacg atgaccagag ctttgggtgtt gggcgtgtt acaaagctga ggttatgt 600  
 ggtgtgtgtt gcaggcggtt ccttggaaaga ccactcctgtt tggagtgtt gggaaagtcc 660  
 aaggcttggaa ggccaatggag gggctgtatgc accccatcc agtggagagg ctgtatgg 720  
 aggcccttctt gatggggagc aaggctatgtt ggtatgtatgc atgaggtaag tgagatggaa 780  
 gtggctgtac gaacagccat tgaggctgag gcaggaggat tgcctaaatgc catagccca 840  
 gtgtcgagac t 851

<210> 40  
 <211> 2204  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3468066CB1

<400> 40  
 gagcagtccct tgctggtccc gcccccgctc ggctgcggc cagggacgc tagtgggtcc 60  
 agtctcttg gcaaccactt gtcctccccc ctcggccctt ttaaccttta gggtgccgg 120  
 gtgcgttgc tctgcgtcc tctcccttt ccccccccccc ttttccacc cccggcgctc 180  
 aggttggctt ggaccggaa cgaagatggc gacttctggc gggcctcg cggagctgg 240  
 gatcggtctt tgcatattcg gcctttaactt actggctatt ttggcattct tctggatata 300  
 ttttcgtaaa taccaaatgc ggcgggaaag tgaagttgtc tccaccataa cagcaatttt 360  
 ttctcttagca attgcactt tcacatcagc actttctacca gtggatataat ttttggtttc 420  
 ttacatgaaa aatcaaaaatg gtacatggat ggactgggtt aatgtatgc tcaagcagaca 480  
 gattgaggac actgtatattt acggttacta tacttttat tctgttatat ttttctgtgt 540  
 gtttcttgcg atcccttttgc tctacttctt tttatggaa aaggatgtt atgataactag 600  
 taaaatgttca caaattaaaaa cggcactcaa gtatattttt ggattttgtt tgatattgtgc 660  
 actgcttctt ttagttgggtt ccttggttcc attgaatgtt cccaaataaca aaaattctac 720  
 agagtggaa aaagtgaatgtt ccctatggat agaacttggaa agtagtcatg gtttagctgc 780  
 attgtcattt tctatcgat tctgtaccc tttatggaaatg ttggcagcta taacttacac 840  
 agcctatggc atgtctgttgc tacctttaaa tctgtataaa ggcacttagaa ggcgtgtt 900  
 tgaacgtttt gaaaacactt aagacattgtt agaagttggaa caacacatcc aacgattaa 960  
 atcaaaaaaggc aaagatgttgc gaccccttgc agcaagggtt aaacgcgcct taaaacaatt 1020  
 tgaagaaaggc ttacgaacac ttaagaagag agagaggat ttgaatttca ttgaaaacag 1080  
 ctgggtggaca aaatttgttgc ggcgtctgttcc tccctgttgc atcgctgttgg gatattttt 1140  
 catcttagtt gcattgttgc ttgtatatttgc tctcttgc tcaaaatgtt ataaagctt 1200  
 tcattcagct ggaatagatt ctggtttcat aatttttgttgc gctaaatgtt gtaatccact 1260  
 gaatatgtt ttcgccttac tacaaacatgt tttccctttt gattatatttca ttataacaat 1320  
 tattattatg tactttatattt ttacttcaat ggcaggaaat gcaaatatttgc gcatatgtt 1380  
 cttttggatt agattatataa aatcagaag aggtagaacc agggcccaag cactcctttt 1440

tctctgcatg	atacttctgc	ttattgtcct	tcacactagc	tacatgattt	atagtcttgc	1500
tccccaatat	gttatgtatg	gaagccaaa	ttacttaata	gagactaata	taacttctga	1560
taatcataaa	ggcaattcaa	ccctttctgt	gccaaagaga	tgtgatgcag	aagctctga	1620
agatcagtgt	actgttaccc	ggacataacct	atcccttcac	aagttctggt	tcttcagtgc	1680
tgcttactat	tttggtaact	gggcctttct	tggggtattt	ttgattggat	taattgtatc	1740
ctgttgtaaa	gggaagaaaat	cgttattga	aggagtagat	gaagattcag	acataagtga	1800
tgatgagccc	tctgtctatt	ctgcttgaca	gccttcgtc	ttaaagggtt	tataatgctg	1860
actgaataatc	tgttatgcat	ttttaaagta	ttaaactaac	attaggattt	gctaactagc	1920
tttcatcaaa	aatggggaca	tggctataag	acaactatata	tttattatat	gtttttctgaa	1980
gtaacatgt	atcatagatt	aaacatttaa	attaccataa	tcatgtctatg	taaataataag	2040
actactgct	ttgtgaggga	atgtttgtc	aaaattttt	cctctaataatgt	ataatagtgt	2100
taaattgatt	aaaaatctc	cagaattaat	atcccccttt	gtcaactttt	gaaaacataa	2160
taaatcatct	gtatctgtgc	cttaggttct	ccaaaaaaaa	aaaa		2204

<210> 41  
<211> 570  
<212> DNA  
<213> *Homo sapiens*

<220>  
<221> misc\_feature  
<223> Incyte ID No: 3592862CB1

```

<400> 41
gccccggggc tcggggagtc ggcgcctatga ccccatcgag gcttccctgg ttgtttagct 60
gggtctcggc cacggcgtgg agagcggcaa gatcaccctt tctgtgtcat tctctgagga 120
aaacaagttc ttctcaagga gaaaagtctg aacttgtcaa acagtccctt aagaagccga 180
agttaccaga aggtcggttt gatgcaccag aggatccca ttttagagaaa gaaccactgg 240
aaaaatttcc agatgtatgtt aatccagtgaa cccaaagaaaa aggtggaccc aggggcccag 300
aaccttacccg atatggagat tgggaacgaa aaggacgctg tattgtatccc taagtgcacat 360
attctttaac ttcaatattt ttttctgtat atgtacatct gaattaactt atttctgtt 420
attttcttc ttatataatcc ttatgtctgt tagtttgtt aatgtgtttt aatataatata 480
tatatatata tatatatata tatatatatgt ggggctttagg aagaaaatat gctgctgtaa 540
atttagaaaaq ggagaccacgc ctgaccaata 570

```

<210> 42  
<211> 802  
<212> DNA  
<213> *Homo sapiens*

<220>  
<221> misc\_feature  
<223> Incyte ID No: 3669422CB1

```

<400> 42
cagggtcaag gtgaagctgg tggtgtctcg aggcgggtga gtgtcatggg ggagcctggg 60
tgggggtcac actggctctc tctagtccca tgtcgtcgtc ctcttcacga tgccctctccc 120
cttcccccaagg gatgtctctg tggagctgcc ttttgttctt atgcacccca agccccacga 180
ccacatcccc ctcgccagac cccagtcagg tgagcacact acccacccca agccctcaga 240
gggaggggct gaagcaggc cagtggagga aaactggccc ttccagcacc caceccccaca 300
cccccttctc cggccccccc agccctctt ccctttccct caccttggaa ctttcccaac 360
caatcccttc acactctctc ccccatcccc ccaagataca cactggaccc tctttgtctg 420
aatgtggca ttaattttt gactgcagct ctgttctcc agccccggcgg tgggtggcaa 480
gctgtgttca taccttaaatt ttcttggaaagg ggacagtgaa aagaggagtg acaggaggga 540
aagggggaga caaaactctt actcttcaacc tcacaccaac accctccctt atcactctt 600
ctgccccccat tccttcaaga ggagacccctt tggggacaag gccgtttttt tttttttttt 660
aaaaagggtt aaaaaaaaaaa gaaggggccca cccctctgaa ggccccccactt cccttggatg ctataatcca 720
atgtggaaat atggcattag ctacaccacc ctgcgccttc ccggatgaa cataccacga 780
actggagatg caagtcctc ag 802

```

<210> 43  
<211> 693  
<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3688740CB1

<400> 43

gttggtttaa	tgggattgtg	gaagagaatg	actccatat	ttggaagttc	tggcacca	60
atcagccacc	gtccaagaac	tgcacacatg	cttacctgtc	tccgtacccc	ttcatgagag	120
gcgagcacaa	ctcgacacct	tatgactctg	cagttattta	ccgtggttc	tgggcagtcc	180
tgtatgctct	ggggtagtt	gctgtagtc	tcgcaagctt	tttgatcatc	tgtgcagccc	240
ccttcgcag	ccatTTTCTC	tacaaagctg	ggggaggctc	atatattgtct	gcagatggaa	300
tttcttcct	ctgttactca	agcctctcaa	agtcccttatt	gtccccagcct	ctgctgtaaa	360
cgttccagc	catcaatgac	atctcactcc	ttcaagccct	tatgccactg	ctgggatgg	420
ccagtcactg	gacctgtcc	acagtgggt	tatattgacc	tttacctcta	cacttgtata	480
taaactcg	ttccatttag	tttcaagaca	cttggaaagca	cagaccaagg	tttacatttgc	540
tgtttaatg	tttttcttgt	aaatgttta	tgcctaaatg	tttctgtact	actcttttttgc	600
ccaaatcctt	attgtttaaa	agtttctctc	ctactataacc	atgccttata	aatattgatt	660
gaatgaatgg	atgaaatgca	tacctgttta	ta	g		693

<210> 44

<211> 1212

<212> DNA

<213> *Homo sapiens*

<220>

<221> misc\_feature

<223> Incyte ID No: 3742589CB1

<400> 44

ccctcgaggc	aacttgcct	tctcaaacat	ggccgccacg	gcccctctgg	aaggaaaccg	60
ctctggggcc	cgcccttgat	ctcggttggg	gggactgggg	atgagagctg	caccgcgcgg	120
gacaagtgc	cggcgccgcc	cgacggagca	gaacagagag	catggagctg	gagaggatcg	180
tcatgtcagc	cctccctggc	ttttgtccaga	cacacccc	ggaggccgac	ctcagtggct	240
tggatgggat	catcttctcc	tatgtcttg	gggtcttgg	ggacactggc	ccctcgggcc	300
catcagagga	gaacttcgat	atggaggctt	tcactgagat	gatggaggcc	tatgtgcctg	360
gtttcgccca	catccccagg	ggcacaatacg	gggacatgtat	gcagaagctc	tcagggcagc	420
tgagcgatgc	caggaacaaa	gagaacctgc	aaccgcagag	ctctgggttc	caaggtcagg	480
tgcccatctc	cccagagccc	ctgcagcggc	ccgaaatgtc	caaagaagag	actaggtctt	540
cggctgtc	tgctcgagac	acccaaagatg	aggcaactgg	cgctgaggag	gagttcttc	600
cagggttgg	tgtactctg	gagggtttcc	ctacccgttgc	ggtggagcac	gcccagtggg	660
tgctggccaa	agctgggggg	gacttggaa	aagctgtgc	gatctgttgc	gaggaaagg	720
aagaggggcc	tgcaagctgg	gagggccccca	acaggacct	gcccagacgc	ctcagaggcc	780
cccaaaagga	tgagctgaag	tccttcatcc	tgcagaagta	catgtgttgc	gatagcgcag	840
aggatcagaa	gattcaccgg	cccatggc	ccaaggaggc	ccccaaagaag	ctgatccgat	900
acatcgacaa	ccaggttagt	agcaccaaa	gggagcgatt	caaagatgt	cggaaacctg	960
aggccgagga	gatgaaggcc	acatacatca	acctcaagcc	agccagaaag	taccgcttcc	1020
attgaggcac	tcggccgact	ctggccgagc	cttcttagct	cagatccca	agggatgcag	1080
gagccctata	ccccatcacaca	ggggccccct	aactctgtc	cccccttctct	actctttgc	1140
tccatagtg	taacctactc	tcggagctgc	ctccatgggc	acagtaaagg	tggcccaagg	1200
aaaaaaaaaa	aa	.				1212